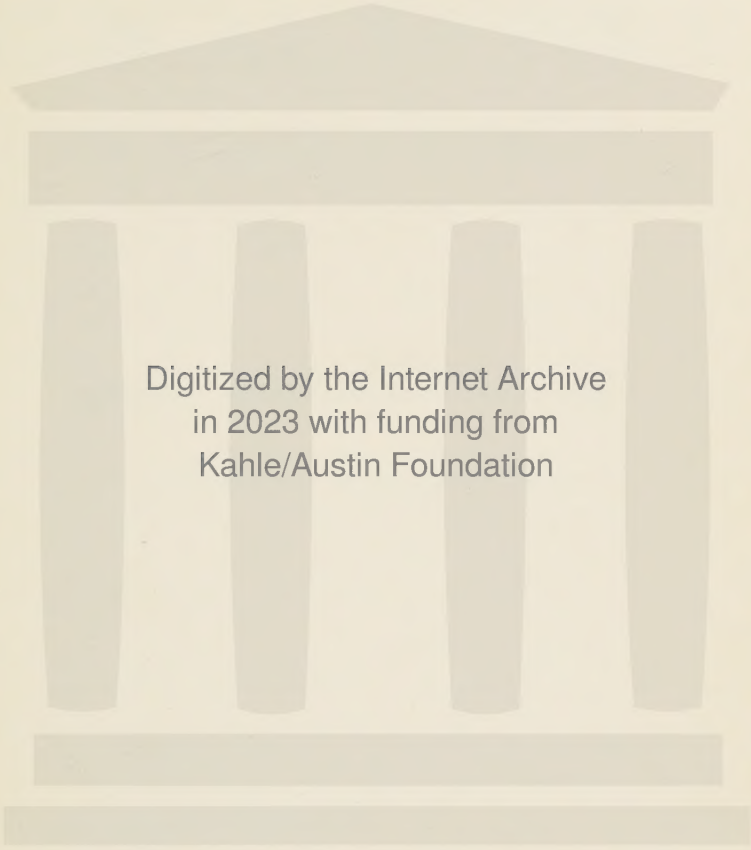
THE MEDICAL CLINICS OF NORTH AMERICA

MODERN CHALLENGES IN MEDICINE

VOLUME 54 – NUMBER 1
JANUARY, 1970

CORNELL UNIVERSITY
MEDICAL COLLEGE
LIBRARY


Digitized by the Internet Archive
in 2023 with funding from
Kahle/Austin Foundation

https://archive.org/details/medicalclinicsof0054unse_s4v5

THE MEDICAL CLINICS OF NORTH AMERICA

MODERN CHALLENGES IN MEDICINE

HANS H. HECHT, M.D., *Guest Editor*

Volume 54 — Number 1

January, 1970

W. B. SAUNDERS COMPANY — Philadelphia • London • Toronto

W. B. Saunders Company: West Washington Square
Philadelphia, Pa. 19105

12 Dyott Street
London, W.C.1

1835 Yonge Street
Toronto 7, Ontario

© 1970 by W. B. Saunders Company. Copyright under the International Copyright Union. All rights reserved. This publication is protected by copyright. No part of it may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying, recording, or otherwise, without written permission from the publisher. Made in the United States of America. Press of W. B. Saunders Company.

The following information is published in accordance with the requirements of the United States Postal Code.

The Medical Clinics of North America is published every other month by W. B. Saunders Company, West Washington Square, Philadelphia, Pennsylvania 19105, at Hampton Road, Cherry Hill, New Jersey 08034. Subscription price is $21.00 per year. Second class postage paid at Cherry Hill, New Jersey 08034.

This issue is Volume 54, Number 1.

The editor of this publication is Albert E. Meier, W. B. Saunders Company, West Washington Square, Philadelphia, Pennsylvania 19105.

Library of Congress catalog card number 17-28505

Contributors to This Issue

GUSTAVO BERMUDEZ, M.D., Research Associate (Instructor), Department of Medicine (Cardiology), University of Chicago, Chicago, Illinois.

RICHARD A. CARLETON, M.D., Director, Section of Cardio-Respiratory Diseases, Department of Medicine, Presbyterian-St. Luke's Hospital; Professor of Medicine, University of Illinois College of Medicine, Chicago, Illinois.

C. HILMON CASTLE, M.D., Associate Professor of Medicine and Professor of Postgraduate Medical Education, University of Utah Medical Center, Salt Lake City, Utah; Coordinator, Intermountain Regional Medical Program.

JAMES J. CASTLES, M.D., Assistant Professor, Department of Medicine, University of Chicago, Chicago, Illinois.

JAMES G. CLARK, M.D., Assistant Attending Physician, Section of Cardio-Respiratory Diseases, Department of Medicine, Presbyterian-St. Luke's Hospital; Instructor in Medicine, University of Illinois College of Medicine, Chicago, Illinois.

ROBERT E. CLEARY, M.D., Assistant Professor, Department of Obstetrics and Gynecology, University of Chicago, Chicago Lying-In Hospital, Chicago, Illinois.

LESLIE J. DeGROOT, M.D., Professor, Department of Medicine, University of Chicago, Chicago, Illinois.

JOHN S. GRAETTINGER, M.D., Chairman, Division of Medicine, Presbyterian-St. Luke's Hospital; Professor of Medicine, University of Illinois College of Medicine, Chicago, Illinois.

ROLF M. GUNNAR, M.S., M.D., Professor of Medicine and Chief, Section of Cardiology, Department of Medicine, University of Illinois College of Medicine; Associate Director, Department of Adult Cardiology, Cook County Hospital, Chicago, Illinois.

HANS H. HECHT, M.D., Blum-Riese Professor of Medicine and Physiology and Past Chairman, Department of Medicine, University of Chicago, Chicago, Illinois.

STANLEY J. HELLER, M.D., U.S. Public Health Service Trainee in Cardiology, Department of Medicine, Presbyterian-St. Luke's Hospital; Research Assistant in Medicine, University of Illinois College of Medicine, Chicago, Illinois.

CYRIL O. HOULE, Ph.D., LL.D., D.H.L., Professor of Education, University of Chicago, Chicago, Illinois.

JOHN E. KASIK, M.D., Ph.D., Associate Professor of Medicine and Pharmacology, University of Chicago, Chicago, Illinois.

ADRIAN I. KATZ, M.D., Assistant Professor of Medicine, University of Chicago, Chicago, Illinois.

SANFORD B. KRANTZ, M.D., Staff Associate, Clinical Pathology Department, National Institutes of Health, Bethesda, Maryland; Assistant Professor of Medicine (Hematology), Department of Medicine, University of Chicago, Chicago, Illinois.

A. M. LAWRENCE, M.D., Ph.D., Associate Professor of Medicine (Endocrinology), Department of Medicine, University of Chicago, Chicago, Illinois.

HENRY S. LOEB, M.D., Assistant Professor of Medicine, University of Illinois College of Medicine; Director, Intensive Care Unit, Cook County Hospital, Chicago, Illinois.

K. G. NAIR, M.D., Ph.D., Associate Professor of Medicine, University of Pennsylvania School of Medicine, Philadelphia, Pennsylvania; formerly Assistant Professor of Medicine and Physiology, University of Chicago, Chicago, Illinois.

OGLESBY PAUL, M.D., Professor of Medicine, Northwestern University Medical School; Chief of the Division of Medicine, Passavant Memorial Hospital, Chicago, Illinois.

RAYMOND D. A. PETERSON, M.D., Professor, LaRabida-University of Chicago Institute and Department of Pediatrics, University of Chicago, Chicago, Illinois.

RAYMOND J. PIETRAS, M.D., Assistant Professor of Medicine, University of Illinois College of Medicine; Director, Cardiac Catheterization Laboratory, University of Illinois Research and Educational Hospitals, Chicago, Illinois.

MURRAY RABINOWITZ, M.D., Professor of Medicine and Biochemistry, Department of Medicine, University of Chicago, Chicago, Illinois.

ERIC REISS, M.D., Chairman, Department of Medicine, Michael Reese Hospital and Medical Center; Professor, Department of Medicine, University of Chicago, Chicago, Illinois.

LEON RESNEKOV, M.D. (Cape Town), M.R.C.P., Associate Professor of Medicine (Cardiology), Department of Medicine, University of Chicago, Chicago, Illinois.

ARTHUR H. RUBENSTEIN, M.D., F.C.P.(S.A.), M.R.C.P.(Lond.), Assistant Professor, Department of Medicine, University of Chicago, Chicago, Illinois.

ULRICH H. RUDOFSKY, Research Associate, Department of Pathology, University of Chicago, Chicago, Illinois.

ANGELO M. SCANU, M.D., Associate Professor of Medicine and Biochemistry, University of Chicago, Chicago, Illinois.

JOHN L. SKOSEY, M.D., Ph.D., Arthritis Foundation Postdoctoral Fellow; Assistant Professor of Medicine, University of Chicago, Chicago, Illinois.

RAYMOND W. STEBLAY, M.D., Assistant Professor and Research Associate, Department of Pathology, University of Chicago, Chicago, Illinois.

DONALD F. STEINER, M.D., Professor, Department of Biochemistry, University of Chicago, Chicago, Illinois.

ALVIN R. TARLOV, M.D., Professor of Medicine and Chairman, Department of Medicine, University of Chicago, Chicago, Illinois.

JOHN S. THOMPSON, M.D., Professor of Medicine, University of Iowa College of Medicine, Iowa City, Iowa; formerly Associate Professor, Department of Medicine, University of Chicago, Chicago, Illinois.

JOHN R. TOBIN, JR., M.S., M.D., Professor of Medicine, Loyola University Stritch School of Medicine, Hines; Director, Department of Adult Cardiology, Cook County Hospital, Chicago, Illinois.

DOROTHY B. WINDHORST, M.D., Assistant Professor, Department of Medicine (Dermatology), University of Chicago, Chicago, Illinois.

RADOVAN ZAK, Ph.D., Assistant Professor of Medicine and Research Associate, Department of Biochemistry, University of Chicago, Chicago, Illinois.

RECENT SYMPOSIA

March, 1969 – *Lahey Clinic*
MEDICAL EMERGENCIES

May, 1969 – *New York University*
REHABILITATION

July, 1969 – *International*
GENETICS IN CLINICAL MEDICINE

September, 1969
PROBLEMS IN REPRODUCTION AND SEXUAL ADJUSTMENT
COMMON EYE DISORDERS AND INJURIES

November, 1969
BASES OF THERAPY

FORTHCOMING SYMPOSIA

March, 1970 – *Nationwide*
DIFFICULT MEDICAL PROBLEMS
MICHAEL G. WOHL, M.D., *Guest Editor*

May, 1970–*Philadelphia*
NEW DEVELOPMENTS IN MEDICINE
HOWARD M. RAWNSLEY, M.D.,
B. S. BLUMBERG, M.D., *Guest Editors*

July, 1970 – *Mayo Clinic*
SOPHISTICATED DIAGNOSTIC PROCEDURES
JOSEPH M. KIELY, M.D., *Guest Editor*

September, 1970
EFFICACY OF ANTIMICROBIAL AND ANTIFUNGAL AGENTS
HOWARD F. CONN, M.D., *Guest Editor*

November, 1970
NUTRITION IN MEDICAL PRACTICE
SEYMOUR LIONEL HALPERN, M.D., *Guest Editor*

Table of Contents

SYMPOSIUM ON MODERN CHALLENGES IN MEDICINE

Foreword 1
Hans H. Hecht, M.D., Guest Editor

I. THE EDUCATIONAL CHALLENGE OF MEDICINE

Introduction 3

To Learn the Future 5
Cyril O. Houle, Ph.D., LL.D., D.H.L.

The Regional Medical Program: The Unicentral (Rural) Region 19
C. Hilmon Castle, M.D.

The Regional Medical Program: The Multicentral (Urban) Area 29
Oglesby Paul, M.D.

A True "Fulltime" Medical School: The University of Chicago System 35
Hans H. Hecht, M.D.; Alvin R. Tarlov, M.D.

II. THE IMMUNOLOGIC RESPONSE—GOOD AND BAD

Introduction 41

TISSUE IMMUNITY AND TISSUE TYPING 43
Raymond D. A. Peterson, M.D.

ALLERGIC REACTIONS TO ANTIBIOTICS 59
John E. Kasik, M.D., Ph.D.; John S. Thompson, M.D.

KIDNEY TRANSPLANTATION: PATIENT SELECTION AND MANAGEMENT 75
Adrian I. Katz, M.D.

NEW FINDINGS REGARDING AUTOIMMUNE FACTORS IN SKIN DISEASE 95
Dorothy B. Windhorst, M.D.; Raymond W. Steblay, M.D.; Ulrich H. Rudofsky; John S. Thompson, M.D.

HEMODYNAMIC OBSERVATIONS IN CARDIAC TRANSPLANTS 107
Richard A. Carleton, M.D.; James G. Clark, M.D.; Stanley J. Heller, M.D.; John S. Graettinger, M.D.

III. METABOLISM AND ENDOCRINOLOGY—NEW CONCEPTS AND METHODS

INTRODUCTION 115

CURRENT CONCEPTS IN MANAGEMENT OF THYROID DISEASE 117
Leslie J. DeGroot, M.D.

PRIMARY HYPERPARATHYROIDISM: A SIMPLIFIED APPROACH TO DIAGNOSIS 131
Eric Reiss, M.D.

SOME ASPECTS OF BONE METABOLISM IN RELATION TO OSTEOPOROSIS 141
John L. Skosey, M.D., Ph.D.

PRACTICAL DIAGNOSTIC AIDS IN THE STUDY OF DYSLIPOPROTEINEMIAS 153
Angelo M. Scanu, M.D.

IV. THE ACTION OF HORMONES

INTRODUCTION 161

Current Status of Oral Contraceptives 163
Robert E. Cleary, M.D.; Rajai M. Dajani, M.D.

Current Status of Erythropoietin 173
Sanford B. Krantz, M.D.

Glucagon in Medicine: New Ideas for an Old Hormone 183
A. M. Lawrence, M.D., Ph.D.

Proinsulin: The Single Chain Precursor of Insulin 191
Arthur H. Rubenstein, M.D., F.C.P.(S.A.), M.R.C.P.(Lond.); Donald F. Steiner, M.D.

The Effect of Insulin on Protein Synthesis in Muscle: An Example of an Approach to the Mechanism of Hormone Action 201
James J. Castles, M.D.

V. NEW TRENDS IN CONCEPTS AND THERAPY OF HEART DISEASE

Introduction 209

Cellular and Subcellular Basis of Cardiac Hypertrophy ... 211
Murray Rabinowitz, M.D.; K. G. Nair, M.D., Ph.D.; Radovan Zak, Ph.D.

The Cellular Action of Digitalis Compounds 221
Hans H. Hecht, M.D.

The Hemodynamic Effects of Myocardial Infarction and Results of Therapy 235
Rolf M. Gunnar, M.S., M.D.; Henry S. Loeb, M.D.; Raymond J. Pietras, M.D.; John R. Tobin, Jr., M.S., M.D.

Electrical Slowing of the Heart and Post-Extrasystolic Potentiation 247
Leon Resnekov, M.D.(Cape Town), M.R.C.P.

The Many Forms of Mitral Incompetence 261
Gustavo Bermudez, M.D.; K. G. Nair, M.D., Ph.D.; Hans H. Hecht, M.D.

Index 275

SYMPOSIUM ON MODERN CHALLENGES IN MEDICINE

Foreword

The impact of new thoughts in medicine reaches a department of a medical school as a wave strikes the jetty in the sea. It is an outpost and a test of strength for new ideas and techniques. Such departments are a testing ground and a screen for all that is new in the understanding of disease, development of new technical skills, and handling of patients. The impatience by which a government wishes to make "new breakthroughs" available quickly to mass-care bodes ill because it tends to circumvent unbiased academic evaluation. When the same forces place overly restrictive fetters on the judicious use of new therapeutic agents, a moral dilemma is placed on academic institutions, which strive best unhampered, unsupervised, and un-site-visited.

The present volume, as others before, presents the distillate of some recent trends in internal medicine as viewed by members of the Chicago faculties. It is by no means all-inclusive, but it points to fields of interest that are likely to be a major concern to the practicing physician in his attempt to understand the patient and his disease. The legislated emphasis on greater interchanges between the community and their physicians on one hand, and the medical schools of a given area on the other makes the inclusion of a section on the education of a physician one of the most pressing modern challenges in medicine. It is fitting that we begin the 1970 issues of the Medical Clinics of North America with such a discussion.

HANS H. HECHT, M.D.
Guest Editor

950 East 59th Street
Chicago, Illinois 60637

I. THE EDUCATIONAL CHALLENGE OF MEDICINE

Introduction

The educational turmoil in medicine of the 60's is likely to be channeled in the direction of much greater emphasis on adequate medical care for all. Thus it will require the most active participation of the practicing internist in programs of continued upgrading of medical skills. Continuing professional education – in all fields – has been greatly stimulated by the searching analysis of Professor Cyril Houle of the University of Chicago who thus sets the tone of the volume. Two examples of Regional Medical Programs – the concrete federal program for better health care through education – are presented in an attempt to bring problems and early results of such efforts to the man who is most concerned but least consulted: the practicing internist. The two sample programs – both called IRMP (Intermountain, and Illinois Regional Medical Programs) involve (1) a plan based on a large rural area with a single medical school at its center, and (2) plans and programs of a complex metropolitan region, several medical schools, and many overlapping committees and considerations. We finally include here an article on the nature of a truly full-time medical system (The University of Chicago) based on a private patient fee-for-service arrangement that in its many applications has rarely been fully understood. The article was written in July 1968 while one of us (H. H. H.) was benefiting from the system as chairman of a major department and as a patient of a major facility. Here again, it is likely that an end of an era is in sight and that new challenges require new approaches.

H. H. H.

To Learn the Future

CYRIL O. HOULE, Ph.D., LL.D., D.H.L.*

When I was young my teachers were the old.
I gave up fire for form till I was cold.
I suffered like a metal being cast.
I went to school to age to learn the past.

Now I am old my teachers are the young.
What can't be molded must be cracked and sprung.
I strain at lessons fit to start a suture.
I go to school to youth to learn the future.

ROBERT FROST

Today everybody talks about continuing professional education—and many people do a lot about it. Furthermore everybody knows what is meant by the term, which has come to be viewed in a rigidly orthodox fashion. It signifies a cluster of special activities undertaken by the leadership groups of a profession to provide sustained lifelong learning for its members. Most of these activities take place in a familiar setting: a lecture amphitheater; a university conference center; a hotel ballroom; or a laboratory, clinic, library, or study. The materials and processes used may also be of a familiar sort—books, journals, lectures, films, slides, manuals, demonstrations, and discussions—but newer devices and techniques are becoming more and more common. Among them are telephone or radio networks, audiotapes, videotapes, closed-circuit or open-circuit television, programmed instructional materials, and computers. Those who take part in any specific activity make two customary judgments about it. How do they feel about the experience itself? More profoundly, how much of the knowledge, skill, or sensitiveness the activity was designed to convey has been absorbed into the understanding and practice of the learners?

The established framework of thought outlined above is based on a collective approach, most commonly that of the single profession as it organizes itself either as a whole or as a group of segments and subspecializations. Ordinarily, too, programs of continuing professional education are defined by a geographic area or service point: the world, the nation, the state or province, the local community (however defined),

*Professor of Education, The University of Chicago.

or the individual institution in which professionals serve their clients. Supporting the collective approach are several ideas: a profession is itself a socially defined entity maintained by both custom and law; the prevailing tradition of education, established by school and university practice, is that learning is a group enterprise; and, most important, a profession has a responsibility to society which it can discharge only if it takes vigorous steps to be sure that all its members keep their performance at the highest possible level.

Valuable though this customary conception has proved to be in practice, few observers of modern continuing professional education are satisfied with the quantity or quality of present activities or with their effect upon practitioners. A critic of efforts to educate practicing school administrators says that "the overall picture, with relatively few exceptions, is one of sporadic activities conducted in rather traditional patterns."[5] His words are echoed by those who report on similar endeavors in other fields of work, and the literature devoted to the subject is filled with criticism (often self-castigation) of inadequate and uninspired effort. Even more ominous is the expression of a growing public hostility toward the several professions because of the alleged incompetence or self-satisfaction of their individual members, faults which better continuing professional education might have helped to prevent.

If this situation is to be remedied, most of the advance will undoubtedly come wholly within the presently accepted framework of thought as the result of the application of greater resources, keener intelligence, and more creative approaches to the task than before. Something of this sort is now happening in medicine with the growth of major federal programs largely devoted to the in-service education of physicians. In engineering, as well, a group of national leaders has developed a "comprehensive concept of continuing engineering studies" broadly similar to the collective approach sketched above. The authors of the report defining this concept go on to say that "glimpsed here is a new advance of learning, of such magnitude as to accommodate and require the best of professional education, the most competent of industrial and government tutors, and the communications network of our well organized engineering societies."[3]

It is the thesis of this paper, however, that the established framework of thought which now guides virtually all continuing professional education needs to be supplemented by other conceptions which can serve as beginning points for both theory and practice, thereby avoiding some of the rigidity which limits present accomplishment. Two such ideas will be presented and developed, the first only briefly, the second at greater length.

A SIMILARITY OF APPROACH

The first idea, implicit in everything said above, is that all professions have marked similarities of approach when they undertake continuing educational programs. The same dominant conception apparently guides all the professions as they turn to the task of educating their members

and, in the process of doing so, they all use essentially the same kinds of facilities, techniques, and ways of thought. This paper is presented in a journal read chiefly by physicians and therefore makes more references to the medical profession than to any other—but the observations it contains are intended to be equally applicable to all professions.*

Yet this similarity of approach is almost entirely ignored in both theory and practice. Teaching and learning lie so near the heart of every profession that the perfection of its practice seems to have no other source than the distinctive knowledge and lore, the special history and belief, and the structures and forms of the profession itself. This mystique is reinforced by the fact that those who consider themselves to be professional educators of adults have not yet established a widespread recognition of their own expertise. A bar association would not insist that its accounts should be audited by a lawyer, nor would a welfare agency require its dental consultation service to be provided by a social worker; in fact, in carrying out such functions, the association would make sure that it had a qualified C.P.A. and the agency that it had a dentist. Yet both association and agency would feel thoroughly competent to undertake their own educational programs and would probably insist that nobody else could undertake them.

The abandonment of the idea that the members of each profession should act as though they alone had any need to learn and that their processes of doing so are wholly unique to them might lead to important consequences. It could and it should result in an exchange of ideas, techniques, and novel approaches which would greatly refresh and broaden current practice. Staff members of associations, university professors, and other people with special assignments to conduct continuing education for the several professions could collaborate, thereby advancing the work of all of them and avoiding that duplication of effort which results from the solemn belief that each one must find its separate road to salvation. In some cases, interprofessional efforts to learn may provide a whole new range of insights. (A special task force made up of a dentist, a pharmacist, a physician, a nurse, and a physical therapist recently concluded, after 3 months of intensive analysis, that collaborative study based on the care of the patient would be desirable and feasible for all the health professions, though such an approach "bore little resemblance to the conventional wisdom about continuing education."[8]) Finally, the field of adult learning would itself be advanced and might perhaps, in time, come to be the basis for a profession or a subprofession. If it is good for a medical specialty group to have its legal matters handled by an attorney rather than by a physician, it would be equally good to have its continuing professional education handled by a person with special training and competence in that field.

INDIVIDUAL EXPERIENCE

A second major idea, often expressed but seldom considered systematically, is that the traditional collective approach should be sup-

*To provide a check on this point, the first draft of the paper was read by specialists on continuing education in seven professions and revised in the light of their suggestions.

plemented by another which focuses upon the individual professional, who is, after all, the ultimate beneficiary of all educational activity. How does learning fit within the pattern of his life? Nobody would expect or want a lifelong curriculum with course following monotonously upon course, but if the term "continuing education" has any meaning at all, it implies sequential experience in which one module of learning, however independently valuable it may be, gains force and direction from the cumulative impact of its integration with other modules.

At present, the most startling and ironic characteristic of continuing education is its discontinuity in the experience of the professional himself. Sequence is seldom planned—at least in the entrepreneurial professions (such as law, medicine, or architecture) or in the organizational professions (such as teaching, engineering, or librarianship). In most of them, the journal, the annual convention, or the regularly scheduled meeting provide as much continuity as is thought necessary, and even in such cases it is only the method of presentation which is continuous, not the messages brought. Other forms of continuing education usually make up an amalgam of bits and pieces of experience. The content of each activity is based on whatever is, for the moment, new, usually some specific area of knowledge or technique which lies at one of the frontiers of rapidly expanding knowledge or practice. The professional learns one element of novelty; later on, perhaps, he learns another; but the second is not designed to be linked to the first in any meaningful way. The growth of knowledge, the refinement of skills, and the deepening of understanding or of ethical sensitivity are achieved, if at all, by what appears to the outsider to be a zigzag and erratic course. Only in the professions which are arranged hierarchically—such as those concerned with business administration, the work of some religious denominations, or the armed services—does the inherent element of administrative control allow for a sequence of learning activities to be arranged so that, as the individual increases his responsibility, he is guided through a series of learning experiences intended to equip him for his progressively broader authority.

But while continuity can be planned by some entity representing the profession or a segment of it, the ultimate test of the existence of sequence lies within the experience of the individual professional. He acquires what he thinks he needs to know from the stream of intellectual nourishment which reaches him from many sources, however rich or meager it may be, and adopts into his practice whatever he wants to use. Like all other aspects of his life, his work has an essential continuity, one phase usually blending into another and new ideas and techniques constantly being absorbed in a gradual fashion.

It is the ideal of every profession, stated or implied in its code of ethics, that each professional should feel a deep and continuing concern that his own education be carried out at a high level throughout a lifetime of practice; therefore, the ultimate responsibility for it can safely lie in his own hands. While all men should share this concern, the professional must feel it with special urgency, for once he achieves his formal status, he is protected in it for life and only rarely can sanctions

or compulsions be brought against him. This freedom carries with it responsibility for effective service, which in turn implies both a rigorous application of what he has already been taught and an assiduous and self-initiated effort to learn whatever helps him improve his service to his clients. The right to educational self-determination is so jealously guarded by most professionals that they reject any effort to interfere with it. Perhaps it is for this reason that the collective framework of continuing professional education has become so traditional; it seems an impertinence to inquire how close a practitioner is to achieving the ideal. But it seems misplaced tenderness to let such an attitude cut off all thought about how a professional actually does pursue his independent inquiries and how he can improve his skill in doing so. If the processes of learning in groups can be studied and improved, why cannot individual study?

The whole subject is so complex that to deal with it fully would require a treatment of massive length. The primary concern, perhaps, is the way by which advancing years change the professional's conception of his role and therefore his learning patterns; but I have already dealt with this subject elsewhere* and now have nothing new to add. Instead of this longitudinal approach, a cross-cutting one will here be used to deal with two questions. What kinds of content should the professional study? How is his conception of continuing education and his practice of it related to his attitude toward innovation in his profession?

CONTENT

In undertaking any given module of learning, the professional usually has some specific goal in mind. He may want to solve a problem, to keep up to date, to equip himself for a new responsibility, to take part in an enjoyed activity, to prepare for a new specialty, or to learn how to practice in a new setting. For each episode, an aim of this sort is sufficient, but if education is to be truly continuing and not merely episodic, some larger framework of purpose is required. It is here suggested that the professional who systematically incorporates learning into his life strikes some balance among four interrelated and overlapping kinds of content, if that term may be broadly conceived to include a full spectrum of knowledge, technique, understanding, and ethical perception. His four endeavors are:

1. *To examine each professional situation with which he deals to find out what he can learn from it.* The essential idea of a profession is that it is a way of focusing advanced bodies of knowledge to deal effectively with specific practical situations. In complex and unusual cases, the need to learn is starkly evident: the professional wants to diagnose and heal the patient, win the case, bring a dissension-wracked institution back to harmonious operation, or build a bridge in a difficult terrain.

*In two papers, "The Lengthened Line" and "A Firstling of His Heart and Hand," both of which appeared in Perspectives of Biology and Medicine, the first in Volume 11, Number 1, Autumn 1967 and the second in Volume 12, Number 4, Summer 1969.

To achieve his goal, he reads, consults other professionals, makes tests and probes, engages specialists, and takes advantage of any other opportunities his resources permit. Whether he succeeds or fails in his immediate task, he learns—and his future practice will be enriched by that fact.

The difficult task is to apply the same kind of intent inquiry required in unusual cases to customary and familiar situations. The difference between the two is only one of degree so far as the benefits of learning are concerned. In every situation in which the professional works, complex factors are present; otherwise his trained talents would not be required. Experience may be a good teacher, but reflection over experience is a better one. A busy, possibly overworked, man may resent the suggestion that he should try to bring a fresh and discerning eye to bear upon even the most routine situation (the treatment of a case of measles, the drafting of a will, the making of a minor executive decision, or the designing of a culvert), but to the extent that he fails to do so, his learning dries up at its freshest and most fruitful source.

2. *To study the literature (either written or orally presented) of the profession itself, either to keep up with new developments or to become aware of previously unconsidered sources.* In most cases, this literature is one of application or of the analysis of practical situations. It explains causes, describes results, suggests alternatives, presents techniques, identifies new approaches, challenges established routines and procedures, and in other ways tells the professional how to cope more effectively than before with the situations he confronts. It takes for granted the fact that he must deal with specific cases and suggests with what degree of certainty he can diagnose, prescribe, design, teach, minister, or perform any other skill required by his profession. Yet the professional literature is always abstract; it deals with only one segment of the full-bodied experience of the professional; therefore it supplements but cannot replace the education which he acquires from his analysis of his own continuing practice. From the literature, he learns the rules; from his experience, the exceptions.

This literature, presented in countless forms, makes up most of the content of continuing professional education and floods in upon the practitioner in increasing torrents which threaten to engulf him. Since it is manifestly impossible to read all the papers, hear all the speakers, or go to all the conferences which bear directly upon his work, the professional may tend to be haphazard in his selection of stimuli or turn away from them altogether except for those brought his way by chance comments, the leafing-through of journals, dicta uttered by authoritative figures at meetings, or the warmly importunate words of salesmen. If he is to avoid this dreary and dulling approach, he must have a plan for learning and give it a high enough priority so that he is not engulfed by the pressures of his daily routine. Such a plan usually has several components: to seek help on particular problem situations; to concentrate on a few journals and read the others sporadically; to select specific lines of investigation and follow them systematically; and, as he has time, to make exploratory forays into other parts of the literature. Somehow

he must strike the balance imposed by his need to practice—he must not read so little that his ignorance does damage to his clients, nor must he read so much that, as was once said of Macaulay, he not only overflowed with learning, but stood in the slop.

3. *To develop a broadening awareness of the disciplines which support his profession.* Every profession has its base in a group of disciplines; in fact, all vocations that achieve professional status put advanced and theoretical knowledge derived from many sources to use in practical situations. Thus the physician knows that the art of healing rests directly upon the biological sciences and that closely related to his practice are many other bodies of knowledge such as psychology, sociology, economics, and statistics. In the course of preservice education, the practitioner of any profession is likely not to get his theoretical knowledge straight, but filtered through the principles of its application. If he does approach a basic discipline directly, he skims its surface, putting down a probe here and there at relevant points. His aim is not to master the knowledge but to see how he can use it. The physician and the physiologist are clearly related to one another but also clearly different, and, as Osler once put it, the doctor must translate the "hieroglyphics [of science] into the plain language of healing."[6]

In the course of a lifetime of practice, the professional is drawn constantly back to a reconsideration of fundamental disciplines and, if he is perceptive, comes in time to understand the unity of all knowledge. The pastor may have taken seminary courses in urban sociology and effective speech, but as soon as he goes on the staff of a slum church, he quickly sees how much more he needs to learn about how cities work and how to use the arts of rhetoric. He also begins to see the relevance of many other matters which were not impressed on his consciousness earlier. Every other professional feels the same kinds of need; more than that, if he is intellectually curious, he is led to try to understand advances and shifts of thought in the disciplines basic to his work. Few people now practicing in the health professions, for example, may find their work influenced by the recent pioneering research in genetics, yet many of them have been curious enough to learn more about it.

Thus the effort to use the years of practice to move outward toward some understanding of basic knowledge is motivated by feelings of both need and interest in constant interplay with one another. Here, too, the professional may profit most if he works out a conscious plan, keeping in touch with the sciences most basic to his work and making occasional explorations into new terrains of related content. The 40 to 50 years of a career pass all too quickly, but they do allow a scope and dimension to continued learning denied in the cramped years of schooling.

4. *To seek new dimensions of knowledge, skill, and sensitiveness by the arduous study of topics not immediately related to his profession.* The first three aims are, in some sense, integrated, moving outward from the immediate situation to the broadest reaches of learning which might influence it. This fourth aim is sharply different. It suggests that the professional should have at least one subject seemingly unrelated to his work in which he is passionately interested and to which he devotes substantial amounts of time.

This aim has clear-cut career values. It often provides a chance to reach new clients or to move in different social circles. More profoundly, it offers relaxation and, paradoxically, the more absorbing the interest, the greater the profit in this respect. "The tired parts of the mind," Winston Churchill once observed, "can be rested and strengthened not merely by rest, but by the use of the other parts. It is not enough merely to switch off the lights which play upon the main and ordinary field of interest; a new field of interest must be illuminated." It is only "when new stars become the lords of the ascendant, that relief, repose, refreshment are afforded."[1] But it was Darwin who pointed out the ultimate value of a broad range of interests to the work of the professional: if he does not possess them, his mind is constricted and stunted, and he can no longer work as creatively as before. Darwin used himself as an example. His work as a scientist gradually came to dominate all his thought and action, driving out his earlier enjoyment of literature, music, and the fine arts. Late in life, he lamented the fact that he had lost his tastes for them. "My mind," he said, "seems to have become a kind of machine for grinding general laws out of large collections of facts, but why this should have caused the atrophy of that part of the brain alone, on which the higher tastes depend, I cannot conceive."[2]

But the broadest value of this fourth aim is not that it aids professional life, but that it provides a corrective for its influence. Too often work dominates life and what a man does becomes a symbol for what he is. While anyone who speaks of chemists, veterinarians, clinical psychologists, or hospital administrators usually intends to refer to the men and women who earn their living in these occupations, the part is often taken for the whole, and the individual is submerged in what is merely one of his many roles. This foreshortened view is destructive of the whole person when applied by others, but is even worse when applied by oneself. In either case, all life comes to be viewed in the increasingly rigid framework of a single profession. Against this danger there is no sure safeguard—but the best preventive throughout life is to have other absorbing interests to serve as a counterbalance to the profession. Then, in the final years, when the practice of the profession must usually cease, the zest for life will continue and an independent dignity and resourcefulness remain.

These four aims of an individually designed and conducted program of continuing professional learning are not put forward here as either moral mandates or guides to success. Nobody has the data to argue successfully at this point that any of them is essential to the maintenance of a career. The young professional can live for quite a while on an accumulated capital of knowledge without replenishing it; the established professional can be maintained by his reputation or the work of his junior associates; and at least a few professionals seem able to rise above all rules. But those who do feel the need to learn usually try to chart some course of action for themselves—and these four aims suggest the broad guides which each may use in developing his distinctive pattern of study.

FIVE GROUPS OF PROFESSIONALS

That pattern varies markedly from one individual to another. Some members of a profession are far readier than others to expose themselves to new ideas and to incorporate innovations into their practice. An effort may be made by the organized profession to maintain that all its members compose a solid phalanx of equally enlightened and forward-looking people, but this fiction is believed only by the unsophisticated. A more experienced observer tends to be judgmental about individual practitioners, designating some as more progressive and forward-looking than others. If this process of comparison could be carried to its ultimate limit, the final result would be at least a rough ordering of everyone within the established boundaries of the profession in terms of his relative willingness to seek out new conceptions and techniques. In that ordering, five groups of people might be differentiated.

The first, here designated the *reinforcers,* stand apart from the other four because, though its members bear the name of the profession, they do not share in its main activity, which is always characterized by disciplined action. A physician is judged by his patients' progress, an architect by the quality of his buildings, and a social worker by his capacity to ameliorate personal and social ills. But the work of the reinforcers lies outside the mainstream of practiced performance, though designed to uphold and strengthen it. They teach, do research, study, organize, administer, regulate, coordinate, and engage in other activities which advance the profession. They work in universities, associations, government offices, foundations, and other places. Collectively the reinforcers make up a group which is high in prestige (however bureaucratic or pedestrian some of its members may be) because of its power to lead the profession and influence the channels of individual advancement of its members. The group is also esteemed because so many of its members stand at the forefront of knowledge, deciding what research shall be fostered, conducting investigations, and disseminating their results by teaching, publication, or in other ways.

Modern sociological investigation suggests that the active members of a profession (as distinguished from the reinforcers) may be ranged along a normal probability curve in terms of their readiness to innovate, as measured, for example, by the relative speed with which they adopt new practices. On this curve, at least four different levels may be distinguished, though not with precision. At the upper end of the continuum are the *innovators,* who consciously seek new knowledge, and many of whom want to be members, even if only peripherally, of the world of scholarship. Their desire to innovate may have to do only with the techniques of their work, but it may also lead them to try to push the frontiers of their profession outward or to make linkages with other realms of knowledge. The *pacesetters* feel the need to be progressive in their practice but are not eager to be the first by whom the new is tried. They wait until a new practice or idea is tested, then adopt it. This recognized judgmental quality of their characters tends to make

them legitimizers of innovation for other people. The *majority adopters* make up the great bulk of the profession. Among them, in an outwardly spreading wave, innovations are gradually diffused, in part because the pacesetters have set a seal of approval on them. The final group are the *laggards*, who learn only what they must know if they are to stay in practice.

Observation suggests (though as yet there has been little research on the matter) that these five groups differ from one another in both the extent and the nature of their continuing learning. Those who adopt practices early are more likely than those who adopt them late to have had exposure to new ideas, which means that they read more journals, go to more meetings, and are more likely to have conscious plans of self-instruction. Most methods of learning can be used by all five groups, but are actually used in varying degrees. It is also true that each group tends to have its distinctive pattern of learning which arises from both the nature of the group and the self-conceptions of its members.

Reinforcers are often so wholly immersed in research and its dissemination that they do not differentiate learning from the other activities of their lives. Paradoxically they may even deny that they are taking part in continuing education, a term they tend to restrict to the traditional forms of learning of professionals engaged in practice. The largest subgroup of reinforcers is usually made up of the faculty members of professional schools, and these men and women spend most of their time in research, reading, discussion, teaching (when well done a form of learning), the holding of clinics, the analysis of cases, and the collective effort to find out how to improve instruction. Periodically they take sabbatical leaves and sometimes they go on collective retreats. They perceive themselves as having a special way of life and do not realize that each of its elements would be regarded by other people as ways of learning. Other reinforcers may be less devoted to inquiry than faculty members, but each subgroup (such as editors, researchers, association executives, or government regulators) tends to have a pattern of life in which learning is central or essential to its work. To some degree, they resemble Sir William Blackstone, author of the *Commentaries on the Laws of England*, of whom it was said that "it was not so much his learning that made the book as it was the book that made him learned."[9]

Innovators, while engaged in practice (often in a highly specialized way), tend to regard themselves as having a marked affinity with reinforcers. They do research, have clear-cut plans of independent learning, seek and cherish part-time teaching positions, belong to groups which have restricted membership and which take pride in being more advanced in their practice than their colleagues, read highly specialized journals, attend invitational seminars, and leave their work occasionally to engage in full-time study. They are also more likely than other practitioners to accept the fourth aim mentioned in the previous section and to seek to master some form of knowledge wholly outside their profession. If so, they usually bring to the study of it that sense of perfectionism which characterizes their customary practice. Like reinforcers, they seldom take part as students in formal programs of continuing

education and may even be contemptuous of them. Innovators are so eager to seek the new that less arduous efforts to learn than their own may seem superficial or even meretricious to them.

Pacesetters look in two directions: toward the innovators and toward the majority adopters. They value opportunities to be exposed to new ideas and techniques, but they also maintain an attitude of some conservatism toward avant-garde practices. Therefore they seek to learn by the methods used by the innovators, though to a less whole-hearted degree. They are likely also to esteem membership in restricted societies, part-time faculty positions, and other similar kinds of learning, though their aim may be not so much to gain knowledge as to prove to themselves and to others that they are part of an inner circle of the profession. And yet the pacesetters are also concerned about the profession itself and feel a strong need to organize, conduct, or take part in its structures and its key functions. Thus the pacesetters strongly support group learning endeavors, giving leadership to such activities as the meetings of professional societies, conventions, exhibits, short courses, lecture series, conferences, and efforts to learn within such institutional settings as schools, welfare agencies, or hospitals. The pacesetters are also the chief sponsors of the newer technologies of learning, such as sound films, closed-circuit color television, videotapes, sound recordings or telephoned digests of research findings, self-scored examinations, computer-assisted instruction, or the use of simulated situations or case studies. The pacesetters often take part in these activities themselves, though they do so with mixed motives; they want to learn but they also want to lend to an activity the support of their presence as acknowledged leaders of the profession.

Majority adopters, since their number is so large, have a broad spread of innovative interests, ranging from those practitioners who are just below the pacesetters to those who are just above the laggards, and sometimes moving upward to one or slipping downward to the other. The majority adopters are the chief participants in organized programs of continuing education, partly because of a sense of obligation and partly because they respond to the example and the appeal of the pacesetters. Ideas and practices are also diffused (more than is sometimes admitted) by the commercial appeals of salesmen, who almost always surround their approach to a professional with some aura of education. They say or imply that they and he both want to serve mankind, and that they can show him straightforwardly how to do so. Majority adopters vary greatly in their acceptance and practice of self-education. They usually pay lip service to the idea, but it is harder to win their attention than their support and harder still to persuade them to participate. They often set themselves apart from innovators and reinforcers, who are, they are likely to think, too extreme in their advocacy of whatever is new or who are likely to live in a world removed from that of practical affairs. The feeling of rejection is reciprocated. J. G. Freymann has provided an example of such separatism:

> . . . among a vocal minority of practicing physicians can be found a thinly veiled anti-intellectualism, preoccupation with medical economics, and resistance to

change. In the academic community can be found islands of intellectual snobbery, startling ignorance of the quality of practice outside the university, and impatience with conservatism.[4]

Such a schism is not unique to medicine, and it is primarily because of this separation between the reinforcers and innovators on one hand and a substantial number of majority adopters on the other that pacesetters are so crucial in any strategy to establish and implement organized continuing education.

Laggards cause deep concern to every profession. It has selected them, trained them, given them license to practice, and must fight to protect the perquisites society grants them. But laggards do not respond properly to their privileged position. Their ideas have hardened before their arteries. Their old skills deteriorate and they do not adopt new ones—or, if they grudgingly do so, it is by a not easily understood process of osmosis or by yielding a bit to pressure. The laggards say (if they defend themselves at all) that education is costly of money and time. They do not realize that ignorance is even more so. The effect of their apathy was once put pungently by Sir George Pickering, Regius Professor of Medicine at Oxford: "Those doctors who completed their formal education a quarter of a century ago and who have not subsequently participated in educational activities are a dead load to the profession and a potential menace to the sick."[7]

To suggest that these five groups exist is not to imply that men are locked universally and forever into a single pattern of innovativeness or of study. Some people may welcome new ideas in one part of their lives and reject them in others; and the growth of responsibility turns many a young innovator into a middle-aged pacesetter. Yet this ordering of individuals into generally defined groups appears to describe at least one facet of reality. If so, the generalized and rather vague idea that everybody ought to learn more than he now knows obscures the existence of specific problem areas, each based on the distinctive learning needs and patterns of a group.

Specifically, the two chief concerns of organized continuing education today in all professions become clearer: how to speed up the learning of the majority adopters and how to reach the laggards with the knowledge they need. The arguments (chiefly among the pacesetters) ordinarily debate whether it is better to use the stick or the carrot. The stick requires the threat of lawsuits or governmental regulation, the fear of adverse publicity, the loss of the right to practice, and the introduction of licenses which must be periodically renewed by examination. The carrot calls for the growth and enrichment of opportunities to learn until they are so pervasive and attractive that they become irresistible. Some greater use than at present of both approaches seems inevitable, though the mixture will vary from profession to profession, as a result of the self-interest of the occupational group itself, of the needs of the community, and of the experience of other professions.

CONCLUDING COMMENTS

This paper has challenged both the isolationism of professions and the restriction of their thought to formal collective systems as they deal

with the nature and needs of continuing education. The central thrust of the argument, however, is not that old systems of thought on this subject should be abandoned, but that they should be enriched by a wider realization of their interdependence with other conceptions. The entire content of this paper illustrates the fact that since all professions embody the same general ideas and procedures in their continuing education endeavors, they can learn together and learn from each other the expertise in adult education which they require, an expertise which is itself becoming recognized as the basis of an independent professional field. And teaching and learning may be examined in either their social or their individual aspects. Collective programs have usually been made the focus of analysis, though they have no other purpose than to aid the individual in his process of learning, both by providing him with knowledge and by helping to foster his spirit of inquiry. The learner himself can equally well be made the center of attention, though he must often turn to organized programs provided for him or which he helps to initiate.

If this more complex view of continuing education proves ultimately to be more valuable than present orthodoxy, the benefits will accrue to the practitioner himself, to his own profession, and to all professions—but most of all to the society which supports and defends not only all of them but the very idea of professionalism.

SUMMARY

In all professions there is a commonality of thought concerning continuing education, yet little recognition of this fact seems evident in current theory and practice. The professions would find it mutually beneficial to engage in collaborative efforts and secure the services of educational specialists in order to create, plan, and develop desired and viable teaching and learning activities. The educational process is both collective and individual in approach; both are essential dimensions and both are interrelated. Each practitioner should try to gain knowledge, technique, understanding, and ethical perception from each problem which his practice brings, from his careful selection of relevant sources in the burgeoning professional literature, from the disciplines which support his profession, and from a mastery of topics apparently unrelated to his profession. The willingness to apply in practice new ideas and innovations is largely dependent upon the practitioner's adopted professional life style, of which there are five: the reinforcer's, the innovator's, the pacesetter's, the majority adopter's, and the laggard's.

REFERENCES

1. Churchill, W. S.: Thoughts and Adventures. London, Thornton Butterworth, Ltd., 1932, p. 297.
2. Darwin, C.: The Autobiography of Charles Darwin: 1809-1882. London, Collins, 1958, pp. 138-139.
3. Engineers Council for Professional Development: Continuing engineering studies. Meeting the challenge through industry, academic institutions, engineering societies and government. A report of the joint advisory committee, 1963-1965, p. 11.

4. Freymann, J. G.: Leadership in American medicine. A matter of personal responsibility. New Eng. J. Med., 270:710-720, 1964.
5. Howsam, R. B.: In-service education of school administrators. Background, present status, and problems. *In* Lynch, P. D., and Blackstone, P. L., eds.: Continuing Education of School Administrators. Albuquerque, New Mexico, University of New Mexico Department of Educational Administration and Foundations, 1966, pp. 16-17.
6. Osler, W.: Aequanimitas. London, Lewis, 1906, p. 31.
7. Pickering, G.: The Challenge to Education. London, Watts, 1967, p. 18.
8. University of Illinois, College of Medicine: Continuing Education for the Health Professions. Report of an Interprofessional Task Force. Chicago, University of Illinois Center for the Study of Medical Education, 1966, p. 31.
9. Warden, L. C.: The Life of Blackstone. Charlottesville, Virginia, Mishie Co., 1938, p. 255.

5835 Kimbark Avenue
Chicago, Illinois 60637

The Regional Medical Program: The Unicentral (Rural) Region

C. HILMON CASTLE, M.D.*

The recommendations in 1964 by the President's Commission on heart disease, cancer and stroke that a network of "regional centers" be established, was replaced by a program in which "regional cooperative arrangements" were to be made among existing health resources. The Act creating Regional Medical Programs provides grants to enable health resources of a region to exercise initiative in identifying and meeting local needs within the major categorical diseases through a broadly defined process. The main reason for this approach emphasizing local initiative was the recognition of political, geographical, and societal diversities within this country, and spokesmen from the nation's health resources testified accordingly during the legislative hearings relating to the Act.

Simply stated, the Regional Medical Program legislation provided our health resources the mechanism for affording the latest advances in medical science for diagnosis, treatment, rehabilitation, and prevention of heart disease, cancer, stroke, and related diseases. Federal support was intended to strengthen volunteer institutions and organizations in their effort to develop local resources to meet local needs and link health resources so that they can be shared and used optimally. A system for sharing facilities and referral of patients is an essential ingredient of regionalization. Medical schools, research institutions, hospitals, voluntary organizations, and other medical institutions, as well as all health professional groups, are encouraged to be included in the "cooperative arrangements." There is no single blueprint to be applied. Advances in knowledge and health needs and the nature of how medical care should be delivered will continuously change; consequently, the arrangements within a region should be flexible and easily changed.

The aim of Regional Medical Programs is to improve patient care, and many believe that this can best be accomplished through the providers of health care. It is believed that continuing education holds the greatest potential for pulling together the health professionals and

*Coordinator, Intermountain Regional Medical Program; Associate Professor of Medicine and Professor of Postgraduate Medical Education, University of Utah Medical Center, Salt Lake City, Utah

bringing about the most rapid and effective solutions to deficiencies in health care. Thus, commitment of the academic and research professionals, as well as the practicing professionals, to the purposes of Regional Medical Programs is necessary for the full potential to be realized.

THE SCOPE OF ACTIVITY

Fifty-five separate Regional Medical Programs have been established to cover the entire country. Forty of these have moved through the organizational and planning stages sufficiently well to establish the process of regionalization and local decision-making and to initiate pilot projects. Others are ready to begin operation, but bureaucracy charged with administering a rapidly growing program with new concepts has slowed program development. Since there were neither organizational frameworks nor experienced staff from which to build this new program at a time when deficiencies in health manpower at all levels were already well documented, development of the first phases has been most difficult. The beginning of Regional Medical Programs was hampered severely because other federally supported health programs emerged almost simultaneously. This has caused confusion, distraction, and antagonism.

In this early stage, Regional Medical Programs have concentrated on organization, relationships, the process of planning, data collection, and decision-making. Even a "tooling-up" period was required before any planning could begin. The requirement for "cooperative arrangements" has provided a not so subtle coercion for medical institutions, organizations, and professional disciplines to talk together about how they can develop more efficient and effective mechanisms for improving patient care. Although enhanced diagnostic and treatment capabilities, and regional "cooperative arrangements" with potential for improved education and patient care within a defined geographic region, have developed in several regions, it is still too early in this new program to measure the impact on heart disease, cancer, stroke, and related diseases.

Three years after the first programs were funded for planning, only about $100 million have been awarded to regional programs. As a result of the slow development, it now appears to be destined for only about $100 million per year for at least the next several years. This amount is significantly less than originally conceived necessary, and too little to have much impact on moving recent scientific advances from the laboratory to the field for application to all the people who would benefit from them.

Types of Programs

By mid-1969 all but 10 of the 55 established Regional Medical Programs were funded for operational pilot projects. Most of the regions have more than 2 million people, and 10 are responsible for over 5 million. Thirty-one programs have as their boundaries state lines, whereas 24 have responded more to patient service areas that either

cut across state lines, such as the intermountain region, or encompass only a portion of a state, such as Albany, New York. The most heavily populated Regional Program area is California, with 19 million people, and the largest in geographic area is Washington/Alaska, with 638,000 square miles. In contrast, the smallest in population is Northern New England, with 600,000 persons, and in area, Metropolitan Washington, D.C., with only 1500 square miles. Some regions are almost entirely urban, such as New York Metropolitan, and others are largely rural, such as Arkansas and Intermountain.

Thirty programs have a university as their headquarters, and 36 use universities as the grantee institution. Only 17 use nonprofit agencies such as the state medical society or newly organized agencies for the specific purpose of developing a Regional Medical Program. Forty-seven of the 55 RMP's have full-time coordinators or directors. Most of them have come from medical schools or universities, while 16 have come from private medical practice and only nine from public agencies, such as health departments. Less than one third of the central planners, coordinators, and administrators of Regional Medical Programs are physicians. A similar proportion of the operational project staff are physicians, but most are only part-time. The major thrust of Regional Medical Programs so far has been continuing education and training in the diseases specified in the Act. Approximately one half of the funds used in regional programs has been used for continuing education, while less than 25 per cent has been used for demonstration of patient care and another 25 per cent for planning, administration, and coordination. Heart disease has been emphasized, and three times as much has been expended for this disease as for stroke and cancer. The amount of funds received by different programs has varied markedly, but on the average the funding per capita per year of the entire Regional Medical Program has been less than 50 cents.

Characteristics of Success

So far, success in the various Regional Medical Programs has been unequal and progress uneven. Analysis of more advanced programs shows some common characteristics which appear to be important to success:

1. The exercise of strong, dynamic, and effective leadership by a full-time coordinator who is able to mobilize the large and varied institutional and health organization leadership to understand and commit themselves to purposes of Regional Medical Programs.

2. Organized commitment of medical centers and medical schools, major community hospitals and their leaders, practicing physicians and officials of the voluntary health agencies, and local health departments obtained through proper involvement of a representative regional advisory group. (Commitment of the health power structure within a region is essential to a program which is expected to have a measurable impact on health care.)

3. Ability to formulate a sound program concept and design within the purposes of Public Law 89-239 and the guidelines provided by the Division of Regional Medical Programs. (The concern is to improve health care through upgrading the skills and services of those who provide care. Also, it is clearly oriented toward specific diseases, despite the hope held by many to have the program otherwise. The program fosters innovation and change, not in the rela-

tionship between physician and patient, but in relationships among providers of health services. Despite the clear mandate to the contrary, some have wanted this program to support the creation of new schools for formal training of health professionals, to develop research centers, and to enhance the capabilities of individual institutions to the exclusion of others in the region.)

4. Effective implementation of pilot projects to give visibility to the program and encourage local identification, acceptance, involvement, and commitment to the purposes of Public Law 89-239.

5. Mechanisms for data collection, evaluation, and feedback to people within the region regarding regional needs, progress, and accomplishments in improvement of health care.

Regions with multiple large urban medical centers have had difficulty in sharing leadership and gaining consensus. Regions without at least one well established medical center with training and research programs in the major diseases and commitment to the purposes of Regional Medical Programs have been equally slow in getting started.

Some programs have elected a prolonged planning phase, while others began early to test ideas in order to gain involvement of the multiple health disciplines. In some, as soon as leadership emerged and organizational involvement began, certain immediate needs and problems readily agreed upon were used for beginning pilot projects. Admittedly, the planning process should be well established, but planning without innovation, demonstration, collection of new information, feedback, and frequent modification of the plan is not the best way to maintain a program when visibility and early accomplishments are essential for continuing support.

THE INTERMOUNTAIN PROGRAM

Planning for the Intermountain Regional Medical Program (IRMP) began officially in February 1966. Planning funds were obtained in July 1966, and the first pilot projects were implemented in April 1967. The intermountain region encompasses 564,000 square miles and includes the state of Utah and parts of Nevada, Idaho, Montana, Wyoming, and Colorado. The total population of this sparsely populated segment of the country is approximately 2.2 million. The area is primarily rural, with a single metropolitan center in Salt Lake City. There are only four communities with over 50,000 people and only 13 with as many as 20,000 to 50,000; there are 30 communities between 5000 and 20,000, and an even larger number with less than 5000 people.

Specialized medical care is not readily available to over half of the people because of the small size of the communities and the distance separating them. This circumstance affects not only the availability of physicians, nurses, and other health manpower, but also the dissemination of new knowledge to those who are providing patient care in the small, scattered towns. Time pressures and distances inhibit both patient consultation and participation in organized efforts at continuing education. Up-to-date facilities for such special services as acute coronary care, radiology, rehabilitation, cancer therapy, and respiratory therapy have been equally hard to provide in this region.

On the other hand, health care resources in the central area and in a few of the larger outlying communities are good and physician-to-population and allied-health-worker-to-population ratios approximate national averages. There are 130 hospitals in the region, but only the University of Utah Medical Center in Salt Lake City, and five other hospitals in Salt Lake City and Ogden, provide house staff training for physicians. Sixteen hospitals in eight communities in the region provide nursing and allied health training.

Leadership for development of the IRMP was provided by the faculty of the University of Utah College of Medicine. The elements of a successful program outlined earlier in this article were obtained and the initial pilot projects implemented in April 1967 emphasized continuing education and involvement of the outlying community hospitals. Subregions were created by forming community planning committees in a dozen major communities and development of learning centers for health professionals in community hospitals. Medical Education Coordinators were recruited and supported in community hospitals, linked with the University Medical Center through a two-way radio communication system and distribution system for television tapes and other teaching aids. Teaching faculty from the University Medical Center, as well as from other national institutions, stimulated practitioners in the Region to become involved in the IRMP.

Realizing that continuing education cannot be effective unless the services needed for excellent patient care are available and accessible to the providers of care, special laboratory services for diagnosis of cardiovascular disease have been provided through a Regional Endocrine Laboratory project. Concentrated training programs for physicians and nurses in coronary care and respiratory therapy have been established in several outlying community hospitals where trained personnel have taken responsibility for training workers in other hospitals in smaller nearby communities. Opportunities for teaching and clinical research have been taken enthusiastically by many of the practicing physicians in the region. In addition to coordination, planning, program administration, and evaluation, the IRMP has specific pilot projects in heart disease, cancer, stroke, and respiratory disease, special laboratory tests previously unavailable in the region, and computer application to on-line patient monitoring, data handling, and multiphasic screening tests.

Education in Heart Disease

The educational program in heart disease concentrates on coronary care as the framework around which a broader, continuing education program directed toward improved patient care is formed. This program is described in more detail to convey the type of projects and activities used in developing a coherent regional program in a sparsely populated area of the country. Our greatest efforts and accomplishments in specific disease categories have been in heart disease, but similar approaches are being used in stroke, cancer, and respiratory disease.

Nurses for Coronary Care Units. When the IRMP was ready to develop pilot projects, our first response was to the demands to train

nurses for the three community hospitals in the intermountain region with coronary care units (CCU). Concentrated courses 3 weeks in length were conducted for small groups of 12 to 15 nurses. Three to six months later, 1-week follow-up refresher courses were given to the same nurses who were trained earlier. By mid-1969, 290 nurses had completed 4 weeks of training in coronary care. Testing for information, skills, attitudes, and performance in simulated and real situations allow them to demonstrate their ability to take primary nursing responsibility for care of patients with acute myocardial infarction.

Although the coronary care training program began in the university medical center, it soon became necessary to decentralize training for nurses in order to use optimally all the resources in the region and to reach those areas where training was needed most. During the first three courses conducted at the medical center, a nurse teaching faculty was developed to conduct courses with similar curricula and evaluation in three of the larger communities outside of Salt Lake City (Reno, Ogden, and Pocatello).

Physicians for Coronary Care Units. Shortly after training in coronary care for nurses was initiated, it became clear that physicians also needed special training, particularly those who were to serve as directors of CCU's. In order to meet the needs in coronary care and continuously train nurses and physicians who treat patients in CCU's in the region, a clinical "core" faculty of physicians and nurses was developed. Physician training consisted of three 3-day courses separated by 3 to 5 months for cardiologists and internists, selected on the basis of their interest and established competence as practitioners and teachers. Thirty-three physicians completed the series of three courses in 1967 and another 32 finished a similar series in 1968.

Evaluation consisted of measuring new information gained by the physicians during their training, and their ability to interpret electrocardiograms and make decisions regarding diagnosis and treatment of patients with acute cardiovascular diseases. In addition, follow-up evaluation determined what they actually do differently for their patients and what new advances they employ in their practice as a result of their training. Of the 65 cardiologists and internists who received training through the IRMP, 54 have been selected to form the clinical "core" faculty, and 25 have already demonstrated their ability and willingness to participate actively as teachers. They, along with the full-time cardiology faculty at the university medical center, form the faculty resource for all training in cardiology in the intermountain region. Similar training in coronary care for general practitioners was initiated in late 1968, and by mid-1969, 50 had received some training and established relationships with the "core" faculty in hospitals with CCU's.

The number of CCU's in the region increased from 3 in 1966 to 12 by 1967, but these were located primarily in the larger communities. Because of the long distances separating many of the smaller communities, it became apparent that CCU's were needed in many other hospitals which served large segments of the region. Strategically located hospitals were assisted in establishing CCU's, so that 90 per cent of the

population in the region would be within 50 miles (or a 1-hour trip by automobile) of a CCU with special monitoring equipment and trained personnel. Community hospital personnel eagerly accepted advice and assistance in developing CCU's. They provided space renovation and most of the monitoring equipment, while IRMP provided plans and training for personnel. In the beginning only 40 CCU's were considered necessary for the entire region, but by mid-1969, 66 CCU's had been established, which brought 90 per cent of the population within 1 hour of a CCU.

Interaction and relationships have developed between the IRMP and all CCU's, and at least one physician and nurse from each have received training in coronary care from the IRMP. Direct on-site assistance in establishing CCU's was provided to only 28 of the community hospitals, but the IRMP served as the catalyst for developing CCU's and source for training of practically all personnel working in them. Despite the accomplishment of developing adequate facilities in coronary care in the intermountain region, in order to maintain optimum care for patients with acute myocardial infarction, continuing education will be needed indefinitely. Educational programs have been started in all CCU's, and the IRMP will provide stimulation and resources to maintain them.

Since the over-all objective of coronary care has been improved patient care, the training programs have been directly integrated with patient care. Evaluation has focused on change in behavior of the provider of care and the relationships and arrangements necessary to sustain learning. In mid-1969 was implemented a regional myocardial infarction data system and a study of effectiveness of CCU's in reducing mortality and morbidity in patients with acute myocardial infarction. This project provides feedback of results and new advances which should be applied to personnel working in CCU's. In addition, this project provides opportunities for members of the "core" faculty in coronary care to engage in clinical investigation and cooperative studies which will add new knowledge in this field.

Computer Applications

On-line computer monitoring of physiologic data in patients with complicated acute myocardial infarctions, as well as in patients recovering from cardiac surgery, has been in operation for over a year in five hospitals in the region through an IRMP project. Service is provided to patients as well as to the physicians and nurses taking care of patients in these hospitals. Decision-making in emergency circumstances and review of important information from the entire course of the patient's illness have been facilitated by the on-line computer system. Use of the computer for screening tests, diagnosis, data collection and analysis, and as an instructional aid have given providers of coronary care a glimpse at what the future holds for computer application in cardiovascular as well as other diseases. The computer project has involved the IRMP from the outset in research and development of new advances which are directly applicable to patient care and education of the pro-

viders of patient care. Also, this project provides a close tie between the IRMP and the research and academic communities.

In addition to the training in coronary care, cardiopulmonary resuscitation (CPR) programs, consisting of instructions, demonstrations, practice of techniques, and evaluation of individual performance, have been conducted for a total of 2000 health professional workers in 41 different communities. Eighty physicians and 84 nurses have received training in Salt Lake City to serve as instructors in outlying distant communities. These instructors organize initial training programs in their own communities, develop a plan of action in their community hospitals, provide continuous in-service training in CPR, and draw on the central IRMP for faculty and teaching aids to help with training personnel throughout the region. In another project, visiting consultants and teachers in cardiology have visited 42 communities to see patients and their physicians in bedside type problem-solving learning sessions. Our manpower resource in heart disease has been strengthened through a clinical cardiology training program in the university medical center, conducted in collaboration with three community hospitals for physicians who wish to improve their knowledge and skills in cardiology. Three physicians have completed a full year of training, while nine others have engaged in 1 to 6 months of formal training.

Our heart disease program represents the largest part of IRMP projects, but similar activities and achievements are emerging in cancer, through use of a regional tumor registry and visiting consultants, and in stroke through a special patient care unit for demonstration of exemplary diagnosis, acute care, and rehabilitation, and through training of the stroke care team, community teaching clinics, and a telephone consultation and information service. Only a relatively small effort has been made in respiratory disease so far, but planning and development of projects in this field and in other disease categories (renal disease and diabetes mellitus) are in progress.

Supervision

All existing projects, as well as those in development, are supervised and coordinated by a central IRMP staff. Considerable effort has been made in developing communication links with community hospitals, educational resources, and staff in distant outlying health facilities. Approximately one fourth of the resources of IRMP are directed toward coordination, administration, planning, data collection, communication, and evaluation.

The development of the IRMP has not been without problems. The major one relates to recruitment, training, and maintenance of a competent professional and administrative staff. In the beginning it was extraordinarily difficult to persuade members of the medical school faculty or practitioners to spend time in planning and developing the IRMP. Even in specific pilot projects directed toward improved patient care through education, the responsibilities, relationships, and rewards still have not been made more desirable than those in traditional and well established positions in the medical school or in most practice arrangements.

All health professionals in the IRMP have received faculty appoint-

ments and have participated as regular members in departments in the university appropriate for their discipline. This has served as an attraction to some, but a deterrent to others, in that salaries are determined primarily by the level of academic appointment. Academic departments have had difficulty coping with appointments and promotions and in providing other rewards for individuals whose contributions are in a new program with purposes somewhat tangential to the departments' traditional responsibilities and commitments. Similar, but much less, difficulty has been encountered in recruiting and maintaining nurses, dentists, allied health workers, and administrators. Without well-defined roles and clear career opportunities, it is natural that the young, vigorous, and potentially productive medical professional worker would be cautious about joining the IRMP full-time. But now that the program is well established with accomplishments and recognition by the medical establishment, the problem of recruitment is becoming noticeably less.

Maintenance of a program responsive to local needs and with sufficient support to insure continuous growth presents new problems. The requirement for setting priorities becomes greater and the task more difficult. Reporting of activities and accomplishments to the Division of Regional Medical Programs (DRMP) and obtaining approval through the current review process for project applications has become long and burdensome. Insufficient funds to cover the entire nation with regional programs has become the major immediate problem for all RMP's. Other challenges facing the well established IRMP include: (1) overcoming bureaucratic rigidity; (2) maintaining clear communication and helpful feedback from the DRMP; and (3) sustaining the flexibility in the IRMP necessary to meet our specific needs.

SUMMARY

The IRMP has responded vigorously to the opportunities provided by Public Law 89-239 to improve health resources and capabilities in heart disease, cancer, stroke, and related diseases. It has been possible to develop an effective RMP with significant accomplishments in less than 2 years. The intermountain region is demonstrating that the RMP mechanism can be a useful instrument for moving medical advances out of the research laboratories and large medical centers and making them readily available to all the people in the region. Problems in initiating this new program have been identified and the major ones overcome. Perhaps the need is greater for "cooperative arrangements" in a large rural region and the problems are more easily managed in a region with only a single major teaching and research medical center. New, and seemingly more difficult, problems related to maintenance of an effective program are emerging to challenge those committed to local initiative in solving regional health care problems. Despite the fact all other RMP's are somewhat different, many of the approaches used in developing the IRMP are applicable to other regions, especially ones largely rural.

REFERENCES

1. Adams, W.: The Illinois Regional Medical Program for heart disease, cancer and stroke. Illinois Med. J., *135*:166-169, 1969.

2. Bank, G., and Mayer, W. D.: Continuing education for the health professions. Missouri Med., *65*:730-733, 1968.
3. Burgess, A. M., Colton, T., and Peterson, O. L.: Categorical programs for heart disease, cancer, and stroke. New Eng. J. Med., *273*:533-537, 1965.
4. Callahan, B.: Regional medical programs taking giant steps. Hospital Progr., *48*:78-83, 1967.
5. Castle, C. H.: Intermountain Program focuses on community hospitals. Hospitals, *42*: 49-51, 1968.
6. Castle, C. H.: The program is regional, the feedback is local; government impact on hospital practice. Hospital Practice, *3*:16-17, 19-20, 24-25, 1968.
7. Castle, C. H.: Regional Medical Programs: Implications for the intermountain area. Rocky Mountain Med. J., *64*:51-55, 1967.
8. Clark, H. T., Jr.: The challenge of the Regional Medical Programs legislation. J. Med. Educ., *41*:344-361, 1966.
9. Hardwicke, H. M.: The Smithville project. Missouri Med., *65*:750-753, 1968.
10. Healey, L. A.: The Washington-Alaska Regional Medical Program. Resident Physician, *14*:58-65, 1968.
11. James, G.: New York physician and the 1970's: The local Regional Medical Program for heart disease, cancer and stroke. New York Med., *24*:194-196, 198, 200, 1968.
12. Jones, F. W.: The medical society and the Regional Medical Program in North Carolina. N. Carolina Med. J., *28*:173-175, 1967.
13. Lewis, C. E.: Local action groups involve communities in Kansas program. Hospitals, *42*:60-62, 1968.
14. Marston, R. Q., and Schmidt, A. M.: Regional Medical Programs—A progress report. Amer. J. Pub. Health, *58*:726-730, 1968.
15. Marston, R. Q., and Yordy, K.: A nation starts a program: Regional Medical Programs, 1965-1966. J. Med. Educ., *42*:17-27, 1967.
16. A national program to conquer heart disease, cancer and stroke. AMA Staff Report. J.A.M.A., *192*:299-301, 1965.
17. Olson, S. W.: Mid-South Regional Medical Program. J. Tenn. Med. Ass., *60*:1072-1077, 1967.
18. Proceedings: Conference on Regional Medical Programs (PHS Publication No. 1682), January 1967.
    a. Farber, S.: The idea, the intent, and the implementation. Pp. 25-28.
    b. Shannon, J. A.: Science and service. Pp. 18-20.
19. Proceedings: Conference Workshop on Regional Medical Programs, Vol. I, (PHS Publication No. 1774), January 1968.
    a. Breslow, L.: Quality and availability of health care for heart disease, cancer, stroke, and related diseases in the future as related to regionalization of health services. Pp. 15-19.
    b. Chapman, C.: Quality and availability of health care for heart disease, cancer, stroke, and related diseases in the future as related to science and service. Pp. 10-14.
20. Proceedings: Conference Workshop on Regional Medical Programs, Vol. II, (PHS Publication No. 1774), January 1968.
    a. Castle, C. H.: Community-centered continuing medical education. Pp. 74-75.
    b. Castle, C. H.: Systems for collection and analysis of clinical data on patients with acute myocardial infarction. Pp. 108-110.
    c. Gilson, J. G., and Castle, C. H.: Development of receptive attitudes toward new ideas. Pp. 25-27.
    d. Haglund, R. F., and Castle, C. H.: An experimental model in organization of a Regional Medical Program. Pp. 144-150.
    e. Smart, C. R.: Cancer training and continuing education and computerized tumor registries, P. 178.
    f. Warner, H., and Budkin, A.: Clinical data collection with a purpose. Pp. 162-166.
21. Russell, J. M.: New federal Regional Medical Programs. New Eng. J. Med., *275*:309-312, 1966.
22. Schmidt, A. M., and Marston, R. Q.: Regional Medical Programs: A view from the federal level. J. Med. Educ., *43*:828-834, 1968.
23. Spiro, H., and DeLuca, V. A., Jr.: The trainee as a teacher in the community hospital. New Eng. J. Med., *276*:903-905, 1967.
24. Turner, G. O.: The community approach to reduction of cardiovascular deaths. Missouri Med., *65*:746-749, 753, 1968.
25. Walker, J. D.: Regional Medical Programs for heart disease, cancer, stroke, and related disease for Kansas. J. Kansas Med. Soc., *68*:162-165, 1967.
26. Wilbur, D. L.: Quality and availability of health care under Regional Medical Programs. J.A.M.A., *203*:143-147, 1968.
27. Wilson, V. E.: Missouri Regional Medical Program. Missouri Med., *65*:719-721, 727, 1968.

University Medical Center
University of Utah
Salt Lake City, Utah 84412

The Regional Medical Program: The Multicentral (Urban) Area

OGLESBY PAUL, M.D.*

In testimony presented before the Subcommittee on Public Health and Welfare of the Committee on Interstate and Foreign Commerce of the House of Representatives in March 1968, Dr. Philip R. Lee described the Regional Medical Program as "already bringing together diverse groups in the health field in an unprecedented fashion and in a manner that results in a consideration of the unfilled health needs of the region. . . ." It would be thought that with the wide awareness of the critical problems facing the low-income, densely populated urban areas, a major and early emphasis of the Regional Medical Program might be directed toward the problems of heart disease, cancer, and stroke in the largest cities. Such was not the case.

For example, when the director of the Division of Regional Medical programs, appearing before the same subcommittee, was asked about forms of organization of specific regions, he cited several examples, none of which included very large metropolitan communities. The 12 operating Regional Medical Programs reported to the Congress at that time were Albany, the Intermountain Region, Kansas, the District of Columbia, Missouri (excluding St. Louis), the Mountain States Region, North Carolina, Rochester (N.Y.), Tennessee Mid-South Region, Washington-Alaska, Western New York, and Wisconsin. On this occasion, the director stated that "the program has tended to go more slowly in the very large, complex urban areas, I think, probably for the same reason some other programs there have gone more slowly." The executive director of the Greater Detroit Area Hospital Council said in a similar vein that "the development of regional medical programs has seemed slow in the inner cities. . . ."

There would indeed seem to be valid reasons why the progress of the Regional Medical Program was relatively lethargic in the largest of our cities, and especially in the ghetto areas. The large number of institutions and agencies existing in the great metropolitan areas presented a formidable problem of representation and coordination. There was some initial apprehension that one or two of the more vigorous and ambitious

*Professor of Medicine, Northwestern University Medical School; Chief of the Division of Medicine, Passavant Memorial Hospital, Chicago, Illinois

participants would dominate the program to the detriment of the weaker members, with a resulting period of maneuver and adjustment. Where the problems were complex, the first step was often a necessary and time-consuming recording of assets and liabilities, which postponed the date of actual operating projects. Further, the urban subregions of greatest need often possessed the least in the way of official and unofficial professional or lay groups prepared to participate; the suburban and rural subregions, with greater affluence, presented a more immediate target for operating projects, because the problems were usually simpler, already defined, and often prepared by prior study for early attack.

Our experiences in Chicago are illustrative of one set of difficulties, having some common features with other major urban areas. It should be stated at the outset that our five medical schools—like many others—had cherished their reputations for research, education, and patient service in that order, and only to a minor degree had interested themselves in community health problems. There had been some limited success in ccoperative arrangements between the schools and nearby hospitals, but little if any interest in arrangements which might extend beyond the immediate metropolitan area.

The concept of any obligation of the medical schools to the as yet nonteaching hospitals and to the medical profession in the Chicago community, in Cook County, or in the state was really nonexistent, with the possible exception of the College of Medicine of the University of Illinois, which recognized some statewide obligation.

The absence of any school of public health in the entire state was also a major barrier. Had such existed, a core of qualified professionals might have been at hand to capitalize more promptly on the concept of a Regional Medical Program.

A brief chronology of the formation of the Illinois Regional Medical Program may be illustrative of the problems encountered in one region including a large urban area. The initiative for formation of an action group to consider a regional medical program came from an urban voluntary health agency, and not the medical schools or organized medicine. It was the Chicago Heart Association which first began discussions with a committee specifically charged to consider the issue. When this committee asked the deans of the five medical schools (all located in the metropolitan Chicago area) to consider working together to seek a joint planning grant, the responses were noncommittal or negative from four and positive from only one. Shortly after this, the Mayor of Chicago forced the issue by appointing a Regional Medical Program Committee, which included all of the deans with the President of the Chicago Board of Health as chairman. Since the Governor of Illinois shortly afterward also appointed a Regional Medical Program Committee, considerable time was spent in combining the Governor's and the Mayor's Committees into a single Committee which was to become the Advisory Committee to the future Illinois Regional Medical Program. The medical schools with their teaching hospitals organized a Coordinating Committee which was to be the actual operating group.

The slow growth and limited success of the Illinois Regional Medical Program to date, particularly in the Chicago area, has been attributable

to many factors, some of which have been listed above. The medical schools and teaching hospitals dominated the Coordinating Committee, yet did not wish to appear to be grasping for power. No other group, however, was in a position to assert more aggressive leadership within the form of organization which existed. Crucial was the reluctance or inability of the schools to assign promptly to the Regional Medical Program qualified faculty members to contribute to planning and operational phases, even though funds were available for such purposes. It was clear that the deans themselves were so burdened with existing problems that they could not give adequate time or thought to the new ones posed by this new and demanding program. Crucial also was the lack of any early decision by the Coordinating Committee to define what type of program was to be encouraged, and to provide any over-all plan as to specific roles of the schools, geographic areas to be emphasized, urban vs. rural approaches, recognition of special needs or opportunities, priorities, etc.

The newness of the Illinois Regional Medical Program, the awkwardness of the relation of the medical schools and teaching hospitals to an undertaking which they had to confront with lukewarm enthusiasm, the lack of specific directions from Washington, and the scarcity of competent professional personnel all contributed to a slow pace. Dr. Carleton B. Chapman, dean of Dartmouth Medical School, pointed to this last problem when he said: "The medical schools themselves have had difficulties in discharging their obligations under the law. Many of them, when the law was passed, had no administrative personnel capable of dealing with these new responsibilities. In some instances programs developing under the law involved commitment of professional personnel that was already fully committed." However, some progress was made. The fact that the five schools worked together harmoniously was an accomplishment in itself; a staff was hired, and under its guidance capable task forces were appointed to survey needs in the area.

Applications for support by the Regional Medical Program were received and reviewed by the task forces, Coordinating Committee, and finally by the Advisory Committee. Three of these involved undertakings clearly outside urban areas, and the remaining five had a variable degree of urban development. None were aimed specifically at health problems concerned with heart disease, stroke, or cancer in the most densely populated and economically disadvantaged areas. As of September 1969, none of these programs were yet operative under the Regional Medical Program.

It is unfortunately true that in most other urban areas, the Regional Medical Program has also been very slow to move. Indeed, the story of slow progress in Illinois and particularly in the Chicago area has been duplicated in most of the large cities. An exception has been in the Watts-Willowbrook district of Los Angeles, funded under the first operational grant to the California Regional Medical Program. The project there involves support for initial planning and development work for the new Charles R. Drew Postgraduate School of Medicine, which will furnish continuing education to physicians and other health personnel from the Watts-Willowbrook area. The school is to be located in the new

Martin Luther King Hospital, currently under construction by Los Angeles County. Trainees, who will be a racially mixed group of interns, residents, and allied health personnel, will be in a program which will emphasize comprehensive continuing care and ambulatory services. Both USC and UCLA have been involved in the staff and faculty appointments.

The New Jersey Regional Medical Program has also created a committee on urban health. On the advice of this group, urban health coordinators for Trenton, Newark, and Hoboken have been secured by the Regional Medical Program, who are responsible for developing the health components of the Model Neighborhood Program. They have been occupied in cooperative arrangements with the local health departments, OEO, voluntary health organizations, hospitals, etc.

The Tennessee Mid-South Regional Medical Program is involved in the support of a long range health evaluation study being carried out in cooperation with the OEO Neighborhood Health Center in North Nashville. The project includes a multitest screening laboratory, and is directed at an indigent black population of 30,000.

The accomplishments by the Regional Medical Program in urban areas may thus be described as at best very modest. This is not to be wondered at in view of the difficulties today with all urban-directed programs. It is not surprising especially in view of the newness of the Regional Medical Program, and the time it took to convince local groups that the initiative in program development was in their hands. Indeed, many could not believe that anyone in Washington would sincerely wish to undertake a federal program involving so much local control and individual planning, and much valuable time was lost in convincing the wary. Today, perhaps the opportunity for success is much greater in the regions that include cities. It is unfortunate that just as this has happened we are faced with serious restrictions in availability of federal funds.

Finally, it would appear to me that a few simple measures might materially assist in pushing the programs along more briskly in the regions and with better chances of achieving useful results. First, the operating committee or corporation—under whatever name—must include enough people interested in and experienced in community health. Men and women with professional degrees and high status in clinical medicine, surgery, or medical administration may be of little or no use in a leadership capacity in Regional Medical Program unless they have had such a background.

Second, the operating group would be well advised to review the local strengths and weaknesses, and to emphasize an over-all but limited plan which is realistic in aims. Too often a wide variety of applications are received and accepted which bear no relation to any well-conceived community or regional plan based upon well-considered priorities. The regional organization should solicit and approve a limited number of applications which have a relevance and applicability to specific aims approved within the region. The very breadth of the possible programs under Regional Medical Program is in itself a source of genuine trouble.

Specific objectives might include an attack on urban problems first, if such exist in the region, or on perhaps one and not all of the three disease groups (heart, stroke, cancer) at the outset, or on education aimed at one group, or on specific cooperative arrangements in one subregion.

Finally, the organization of the region should involve the Advisory Committee required by the law in the program in a meaningful fashion. Such a committee is the nearest thing to the voice of the community in the Regional Medical Program. If it is not involved in policy-making or at least discussion of policy, it becomes a disinterested and useless body. The operating group will better assure its own success by making the Advisory Committee an obligatory full member of the team.

The genius of the Regional Medical Program lies in the two words "cooperative arrangements." Properly realized within both the structure of the organization regionally and in its program, the Regional Medical Program has the potential for being one of the most invigorating and strengthening forces in our national health scene.

ACKNOWLEDGMENT

A portion of the above material may be located in Excerpts of Testimony and Materials presented on March 26 and 27, 1968, Re: Sections of H.R. 15758 relative to Regional Medical Programs before the Subcommittee on Public Health and Welfare of the Committee on Interstate and Foreign Commerce, House of Representatives, Ninetieth Congress, Second Session.

303 E. Superior Street
Chicago, Illinois 60611

A True "Fulltime" Medical School: The University of Chicago System

HANS H. HECHT, M.D.*

ALVIN R. TARLOV, M.D.**

A characteristic of a live medical school is the constant major and minor adjustments of its curriculum. The recent educational turmoil in almost all major universities has had and will have its impact on the curriculum of a medical school. In this context, the somewhat unique and generally poorly understood educational structure of the clinical departments of the Biological Sciences Division of the University of Chicago might be of interest.

The beginning of this medical school had a tortuous history. Upon the framework of a universal undergraduate college founded in 1892, the growth of graduate schools proceeded fairly smoothly under the first president of the university, William Rainey Harper. Several graduate divisons were in full operation when in 1899 degree requirements were adjusted to allow students to take premedical course work. With proper qualifications, a student was permitted to take his first 2 years in medicine or was allowed to transfer into the university for a 2-year program which, the faculty was assured, consisted "almost entirely of courses in pure science." While the practical problems of establishing a graduate school for social science, law, or technology at a new university were relatively minor, the clinical facilities necessary to conduct a medical school required either an entirely new foundation or an "affiliation" or "union" with an existing medical school. Rush Medical College, a proprietary medical school on the West Side of Chicago, expressed an interest in a cooperative venture and eventual union with the University of Chicago. It transferred its first 2 years of instruction to the University of Chicago, raised admission requirements, and let the entrance examinations be given by the university faculty. These attempts at fusion were actively encouraged by Harper, and by E. Fletcher Ingals and Frank Billings (both "Professors" at Rush).

From the Department of Medicine, The University of Chicago

*Blum-Riese Professor of Medicine and Physiology and Past-Chairman, Department of Medicine

**Professor of Medicine and Chairman, Department of Medicine

Around 1900, the establishment of a 4-year medical school at the University of Chicago seemed a near reality. The financial backer of the university, John D. Rockefeller, Sr., and his advisors, notably Frederick T. Gates, however, urged great caution in dealing with proprietary schools in the pre-Flexner era, particularly with Rush, which was financially unstable and unable to raise a necessary "dowry." The Rockefeller Foundation was at the time ready to support and endow a school or Institute devoted to medical research, while Harper and his trustees insisted that teaching must be an integral function of any graduate school at Chicago. As if in answer to this, the Rockefeller Foundation established in 1901 the Rockefeller Institute for Medical Research in New York (ironically now also a "University"). Following this sour note, no further moves were made toward establishing a medical school during Harper's lifetime.

In 1923, after a more successful fund-raising campaign during which Frank Billings raised over $1 million from his own family for the construction of hospital facilities, a committee of the Senate of the University reported to President E. Burton:[3]

> The aim of the University of Chicago Medical School should not be primarily to increase the number of practitioners... Further progress in medicine depends upon the advancement of medical knowledge, and it is believed that the University of Chicago is in a peculiarly favorable position for promoting research and training investigators in the medical sciences. The strength of the basic departments of physics, chemistry, and biology, and the research spirit with which these departments are imbued, the more or less complete organization of the preclinical departments which already exists, the unusual opportunity for developing hospitals, research institutions, clinical departments, and medical libraries in close contiguity to the present scientific laboratories, the strategic position of Chicago and its probable influence upon medical development throughout the Middle West and the South, are all reasons for believing that a University of Chicago Medical School should be founded, having for its chief aim the advancement of the medical sciences.

With the consent of the Rush Medical College, affairs now moved quickly forward, and instructions in the clinical departments of the University of Chicago began in the fall of 1927. The "union" was not absolute, since clinical teaching continued at Rush for some time, and 2-year graduate students from other universities were also accepted there. (Its plant, including the library, eventually became part of the clinical complexes of the University of Illinois and is still a visible component of the Presbyterian-St. Luke's Hospital in Chicago.) At the University of Chicago, the first clinical complex (Albert Merritt Billings Hospital) rose next to the oldest building on the campus (Cobb Hall) and adjacent to the General University (Harper) Library.

This brief historical account is of interest, since it reflects the attitude toward a clinical setting for medical students at this university, an attitude that has remained essentially unchanged for 70 years. Yet it appears that because of the need for expanded community services, planned affiliations with outlying hospitals, and participation in the Illinois Regional Medical Program, fundamental changes in this general structure may be around the corner: an end of a particularly flavorful era appears in sight.

In keeping with this spirit of an integrated academic community of scholars, two important restrictions were placed early on the operation of the clinical departments which, while logical from the university's point of view, are unique among medical schools in this country. They have remained unchanged and are still in force today:

1. A professor in an academic department of surgery or obstetrics or any other clinical unit is equal in all respects to a professor in any other academic department such as history, Assyrian languages, or atomic physics. He is "full time" and has no outside income derived from his clinical skills. There is, therefore, no provision for a part-time clinical faculty.

2. The university has been unwilling to become financially involved in hospital or clinic service operations. The hospital had to remain solvent on its own feet. While it soon built up a reserve for special free-care admissions, the clinical service had to rely for its teaching, research, and demonstrations almost exclusively on private and third party pay cases. This has not presented major difficulties. Then, as now, the hospital collects all fees, and the physician—regardless of his skills, reputation, or the number of patients who seek him out personally—receives only his monthly paycheck. This corporate system of medicine and the low fees charged for services in the early days of the medical school was anathema to the AMA for many years, and to this day only a few members of the Chicago faculty have held elective offices in county or state medical societies.

While all patients are in essence the responsibility of the department as a whole, in many instances a faculty member began to select a special category of patients. He often becomes a specialist from the point of view of teaching in the rapidly advancing fields of medicine. His research requires a scholarly approach to a limited scope of diseases. The system involved then was one in which most members of the department selected a specialty area, joined a group responsible for care and scholarship in this area, and in the main concentrated their practice (both consulting and actual continuing care) in this "subspecialty." Each man appreciates even today that in the natural selection of clinical material, a number of complications or associated conditions will be encountered which require him to remain fairly proficient in other areas of internal medicine as well.

Frequent (and free) consultation was always encouraged and practiced. It provides broad faculty participation in the care of patients and widens the scope of the special patient material available to the faculty.

The system has led to certain consequences, some good, some questionable. Using the Department of Internal Medicine as the example (other clinical departments do not differ strikingly) the following observations can be made.

1. The absence of a part-time clinical faculty to assist in teaching students, for example in an outpatient department, and the use of private patients, makes for a strong commitment of the faculty to patient care—in other words, the teacher remains a finely honed clinician.

Whether he wants to or not, he finds himself at times in competition with practicing physicians of the community, particularly if the referral clinic aspect is not always strictly adhered to.

2. Because of the absence of a part-time group to share the load, the full-time faculty is very large, and outside contacts with community physicians tend to be small. A good part-time faculty elsewhere usually attends and often adds to staff rounds and seminars. This is not a common experience at Chicago, a situation that, in spite of inherent difficulties, the present faculty is attempting to correct.

3. The Department of Medicine consists of 130 full-time staff members, including residents and interns but excluding fellows and trainees. To remain a coherent department today requires a faculty with interests in various aspects of research as well as in the many subspecialties. This large faculty ranges in skill and interests from the purely research-oriented physician who sees few if any patients, to the individuals (rarae aves) who are capable of carrying out basic physiological or biochemical research but are able to maintain their clinical aptitudes, to individuals in whom teaching and patient care is a form of clinical investigation that goes hand in hand with laboratory commitments, but where the latter might at times be subordinated to the former. All members of the faculty have their own research laboratories and in general provide their own training and research funds.

4. The size of this truly full-time department is large not only because of its need for faculty, but because such a faculty is strongly academic and research-oriented. All sections and most senior staff require secretaries, and the central office of the department needs executive officers, assistants, and accountants to handle the large volume of grant applications alone: during the fiscal year 1966-1967, the Department of Medicine administered 11 training grants and 137 research grants from various agencies to a total of $3,355,259. The faculty payroll of the department, including residents but excluding interns, fellows, and trainees, was $1,136,047 for the same period. This requires a supporting staff of 28 secretaries, five full-time accountants, and next to the chairman, a staff secretary and an executive assistant.

5. Two or three drawbacks became apparent fairly early. In order to be manageable at all, the subspecialty services ("sections") had to assume considerable responsibilities, including the supervision of subspecialty inpatient services ("subspecialty wards"). In the hands of a strong organizer, they have tended to become semi-autonomous units as the work load, the research involvement, and the staff increased. As stated above, it must be understood that a subspecialty service never confined itself to a subspecialty exclusively. At a given time, perhaps only half of the hospitalized patients have an illness related to the subspecialty on whose service they are admitted. With the developing autonomy, the inpatient service became more of a private clinic, with the house staff occasionally given less responsibility than desirable in a university hospital.

In the early 1960's, while other institutions began to move toward strengthening subspecialty services, the threat of an overspecialized system for full-time faculty departments became apparent, and at

Chicago, a move began in the opposite direction to restore a needed balance. This included three important steps:

1. All medical inpatient services strengthened and expanded the responsibilities and teaching assignments of the attending physicians. For a given period, the ultimate patient responsibility lies with the attending physician and his house staff. During the period of inpatient care, the patient belongs to the service, and on discharge is returned to "his" original physician (who may be a member of the section but not an attending physician at the time).

2. The creation of a general medical service vis-a-vis the "sections." There were always some faculty members who did not feel the need for a highly specialized area of clinical practice to advance their academic aspirations. They became general internists, a direction that is now strongly encouraged. They now have a new home and renewed responsibilities, since this service in particular aims to develop the primary interests and skills of the residency staff who admit here all medical cases from the admissions and emergency rooms needing hospitalization and not specifically called in by the subspecialty services. This service now consists of 55 beds, or approximately one fourth of all hospital beds assigned to the Department of Medicine. It is, therefore, the largest individual inpatient service of the department, and since admission is accomplished by direct transfer from the emergency room, it has become the most popular rotation for students and house staff alike. This service restores a balance and serves as a spawning ground for training in the general clinical aspects of internal medicine. It has been of great interest that this return to the old-fashioned general ward system has been so successful. This is presumably so because it is now supplemented by equally well qualified subspeciality services. These do not compete with general medicine, but provide a stimulating foil both from the viewpoint of patient care and teaching. Many of the "subspecialists" serve as attending physicians on the general medical service by invitation. The system provides then an unusually well-balanced training program in graduate and postgraduate internal medicine.

3. The third unifying feature consists of a daily brief morning report of all residents assigned to inpatient services to the chief resident and to the chariman of the department. These 30-minute sessions, where the events of the preceding 24 hours (admissions, discharges, deaths, complications) are briefly presented, give everyone, including the chairman, an overview of the activities of the department and of the individual sections. In addition, it establishes a unifying link between the house staff themselves and provides for a daily personal contact with the departmental chairman.

To remain balanced and flexible and to prevent stagnation, such a system is under constant surveillance by the faculty. Provincialism and inbreeding, favoring a "status quo," which is common in large institutions, is avoided if the "system" is regarded as malleable to suit the needs of the present.

In conclusion, the clinical setting at the University of Chicago is large, expensive, and easily calcified. It, however, has a sound basis: its clinical responsibilities and teaching are carried out by a completely

full-time staff, its full interdigitation with and proximity to the university proper provides a fertile ground for basic research in biology, in contrast to the type of "clinical investigation" in vogue some years ago. The recent emphasis on a department as a whole in contrast to the idea of a federation of subspecialties has expanded its basis and has provided an unusually broad setting which should prepare it for the unusual educational challenges that all medical schools face today.

AUTHORS' NOTE

This article was prepared and circulated in 1968. More recently, we note that Dr. H. Ricketts, emeritus Professor of the University of Chicago, has given his personal views and experiences with the system in an article which appeared in the Journal of the American Medical Association, *208*:2069, 1969.

REFERENCES

1. Levi, E. H.: The university and the modern condition. An address to the University of Chicago Citizens Board Luncheon, November 16, 1967, published by the University of Chicago, 5801 South Ellis Avenue, Chicago.
2. Storr, R. J.: Harper's University, The Beginnings. Chicago, University of Chicago Press, 1966.
3. Veith, I., and McLean, F. C.: The University of Chicago Clinics and Clinical Departments, 1927-1952. Chicago, University of Chicago Press, 1952.

950 East 59th Street
Chicago, Illinois 60637

II. THE IMMUNOLOGIC RESPONSE – GOOD AND BAD

Introduction

There are few potentially incendiary areas in medicine which have had a more striking impact on almost all subspecialties of medicine than the science of tissue reactions viewed as a cellular response to the environment. Of interest primarily to those concerned with acute or delayed sensitivity reactions and the allergic response, immunology as a science has in the past been relegated to a subsidiary field of the general area of infectious disease and host response. The growth of cellular biology and the development of special techniques designed to label intracellular components and to separate protein fractions have made immunology in its widest sense the science which defines individuality. It has become of urgent concern to clinical and investigative problems in neurology, dermatology, hematology, cardiology, gastroenterology, and oncology, to name just a few subdisciplines of medicine.

This section properly occupies the center of the stage in the present volume. It begins with a discussion of general problems of tissue immunity and tissue typing by R. Peterson, followed by an extensive survey of the present state, indications, and management of kidney transplants. Given a successful transplantation, the integrated function of the new "foreign" organ in contrast to the simple, relatively nonreactive transplantation of organ parts (cornea, heart valves) has not been given much attention. This is the subject of a brief summary by R. Carleton based on a carefully studied example of a human cardiac transplant.

Organ transplantations, while dramatic and great for fund raising, occupy only a small part of the field. More widely applicable information is presented in a few introductory articles dealing with immunological considerations of the nuclear material of the skin as a target organ for a variety of systemic disorders, and tissue reactions to antibiotics. A discussion of the disordered cell response in malignant growth could not be readied in time and remains therefore a significant omission. This section is nevertheless intended to open the door to a further understanding of this all-pervasive field.

H. H. H.

Tissue Immunity and Tissue Typing

RAYMOND D. A. PETERSON, M.D.*

Tissue immunity is an immune response to foreign tissue antigens. Thus defined, it seems but a colorless biological phenomenon, one that could interest only some ivory tower scientist. A closer look at tissue immunity, however, quickly reveals that it is far from drab. Tissue immunity thwarts organ transplantation, and for that reason it is the subject of much concern today. Tissue immunity seems to protect us from some types of cancer, and thus it must be understood not only by physicians but by research scientists concerned with methods of controlling malignant growth. Tissue immunity protects us from infectious agents such as yeasts and certain protozoa. Finally, tissue immunity appears to be responsible for tissue damage in some autoimmune and infectious diseases. Tissue immunity thus deserves a clear presentation in order that this fundamental, fascinating phenomenon be widely understood and appreciated.

Tissue typing is a laboratory technique designed to define the antigenic composition of tissue surfaces. Tissue typing is thus a technique of interest in both the practical and the research aspects of tissue immunity. On the practical side, tissue typing can improve organ transplantation by facilitating donor-recipient pair selection; on the research side, it can provide insight into many fundamental biological processes—for example, the events surrounding the transformation of a normal to a malignant cell.

It is the purpose of this paper to review the subject of tissue immunity and tissue typing. We will begin with a look at the historic aspects of tissue immunity—especially the observations that led to an appreciation of individuality and of the role of the immune system in the rejection of foreign tissue transplants. Next the chemistry of the factors responsible for this individuality and the methods used for their detection will be discussed. The immunologic mechanisms responsible for tissue rejection will be presented and, finally, the clinical implications

---

From the Variety Club Research Center of La Rabida-University of Chicago Institute and the Department of Pediatrics, University of Chicago

*Professor, La Rabida-University of Chicago Institute and Department of Pediatrics, University of Chicago

Supported by Sprague Memorial Institutional Grant, American Cancer Society, American Heart Association, and NIH Grant AI-08771-01.

of the entire subject will be touched upon as a prelude to the other sections in this volume.

INDIVIDUALITY

The literature of the past is speckled with fantasies and case reports indicating that man did not really appreciate how truly *individual* a person was. Reports of the actual or imagined transplantation of parts from animals to man or man to man are common.[16] Although some of the reports imply success, it seems fair to assume that those optimistic reports were fallacious. There were persons centuries ago who recognized the uniqueness of individuals, but it was not until the late 1880's that individuality as it pertained to tissue transplantation came under systematic study. Two experimental approaches were taken that eventually fused into the scientific discipline known as transplantation biology. One was concerned with the transplantability of tumors among various strains of mice and the other with the transplantability of normal tissues, particularly skin.

Most modern knowledge of individuality comes from experimental studies performed in the mouse. Mammalian genetics really became an experimental science with the development of inbred strains of mice. This development, motivated in large part by curiosity about transplantation, permitted the burst of tumor and skin transplantation experiments that is the basis of our present knowledge of tissue immunity.

An early observation of Loeb's set the tone for many future experiments.[36] He noted that a mammary carcinoma grew well in a partially inbred strain of Japanese waltzing mice but would not grow when transplanted to an unrelated mouse strain. Further studies confirmed Loeb's initial suspicion that the "susceptibility factors" determining the transplantability of tumors were dependent upon the genetic information present in the tumor cells. In 1914 Little formulated a theory to explain the inheritance pattern of the susceptibility factors.[35] He proposed that acceptance of a tumor graft depended upon the graft and host possessing a number of susceptibility factors in common, and that each of these was determined by an independent gene. Further study confirmed the theory. Snell, in an extensive systematic series of skin transplantation experiments, proved that the same situation existed with nontumor tissue.[62-64] He suggested that the genes which determine the fate of transplants be termed "histocompatibility" genes—a term that has become standard in transplantation biology jargon.

Thus it was that individuality became recognized. Animals, mice or men, inherit certain characteristics that permit or thwart transplantation of tumors or normal tissues. These characteristics are inherited via multiple histocompatibility genes.

IMMUNITY AND TISSUE TRANSPLANTS

It was many years before the immune system was firmly defined as the mechanism responsible for the rejection of a foreign tissue trans-

plant. The definitive studies of Medawar, conducted in the 1940's, clearly demonstrated that skin transplants exchanged between unrelated animals would be rejected in accordance with the rules of immunology.[38] For example, the first time grafts were exchanged between two unrelated animals, A and B, the grafts became vascularized and appeared healthy for several days, but then underwent rejection. This slow rejection following initial acceptance was termed a "first set" rejection pattern. Repeat exchange of skin grafts between A and B resulted in accelerated rejection reaction, often before any vascularization occurred. This "second set" rejection pattern was evidence of an altered reactivity conditioned by a previous exposure; a phenomenon we consider the sign of immunity. Skin from animal C, unrelated to either A or B, would be rejected by A and B with a first set pattern, indicating that the phenomenon met the other criterion of immunity—namely, specificity. These types of experiments, generally performed in various inbred strains of mice, established the central role of immunity in the recognition of foreign tissue factors and in the ultimate rejection of such tissues from the body.[2, 10, 39]

The realization that the inherited susceptibility factors were important because they elicited an immune response when transplanted to an animal who did not have the same factors on his own cells brought transplantation biology into the field of immunology. Immunology, a relatively new discipline itself, having developed in the late nineteenth century, brought useful concepts and techniques to bear on the problems of transplantation biology. The resulting fusion of disciplines deepened knowledge of both fields of study and laid the foundation for the broader and extremely productive study known as immunobiology.

TRANSPLANTATION ANTIGENS

Susceptibility factors are critical to tissue transplantation because they are antigenic. Therefore, the terms transplantation or histocompatibility antigens have replaced the previous designation. Perhaps one slight qualification should be made, however, before assigning all susceptibility factors to the role of antigens. There is evidence that cells do not survive as well when placed in contact with foreign cells and that this phenomenon is independent of the immune system. This phenomenon is described in the literature under the heading "allogeneic inhibition."[29, 30] It will not be further discussed here, since it does not appear to be a major mechanism involved in tissue transplant rejection.

Transplantation antigens are synthesized by cells according to the now well described processes of protein synthesis. Their amino acid sequences are determined by the nucleotide sequences of the DNA. They are present on the surfaces of most cells of an animal, although their concentration on any one cell type may vary considerably among species. These cellular antigens are the hallmark of individuality. Their chemistry, the techniques for their detection, the inheritance pattern of the genes that code for them, the manner in which they elicit an immune

response, the nature of the immune response they elicit, and the biologic significance of all this will now be sequentially presented.

Chemistry of Transplantation Antigens

The purification and characterization of transplantation antigens has occupied the attention of several laboratories for the past 5 years. It is now possible to separate these antigens from cell membranes, solubilize them, and perform preliminary studies aimed at determining their structure.[20, 32, 33, 47, 69]

Transplant antigens appear to be bound to cell membranes. They can be released either by autolytic processes or by papain digestion. The soluble products have been preliminarily identified as glycoproteins, with a molecular weight of 50,000. At the present stage of purification the major transplantation antigens of both mouse and human cells appear to be similar types of macromolecules. Eventually, it will no doubt be possible to isolate individual antigens, determine their amino acid sequences, and then really compare the structure of individual antigens from various species. This field of study is just beginning to produce results.

Detection of Transplantation Antigens: Tissue Typing

For both practical and theoretical reasons, tissue transplants are no longer the most useful method of testing for transplantation antigens. Other techniques are now available that provide much more definitive information about these antigens. The term tissue typing has been applied to these latter techniques. It is a somewhat presumptuous term, in that it implies more discrimination than is actually possible.

Now that kidney transplantation has become an effective part of our clinical armamentarium, the proper selection of donor-recipient combinations has become the major problem for transplant teams around the world. This situation has brought the issue of tissue typing into clinical medicine: the analysis of tissue typing data is the crux of most selection decisions. An understanding of the general principles involved in tissue typing is, therefore, essential. Perhaps even more important is an appreciation of the limitations of present day tissue typing capabilities.

IN VIVO METHODS. *Lymphocyte transfer test.* This test involves the transfer of blood lymphocytes from one animal to an intradermal location in another. The appearance of an area of redness and induration implies that the transplanted cells and the host are disparate regarding important transplantation antigens.[14, 80] For several reasons this test has not proved useful in man.[25]

Neutral third party test. In this test an innocent bystander, generally an irradiated hamster, is used as a site for the interaction of cells from two other sources.[67] For example, leukocytes from persons A and B are mixed in a syringe and injected intradermally into the "neutralized" hamster. An inflammation at the site implies tissue incompatibility.

Graft versus host assay. Next to the skin graft experiments, probably the graft versus host assay has contributed more to clarifying

individuality than any other test.[58] Lymphocytes from one animal are injected into either a newborn animal or an adult animal who for some reason is unable to reject the donor cells. Several days later the recipient animal is sacrificed and its spleen weight analyzed. A large spleen implies that the donor cells reacted against the host spleen cells. There are also other methods of setting up the experimental situation, but the principle is the same in each case. This test has permitted the study of the inheritance pattern of transplantation antigens in a number of experimental animals. The phenomenon of a graft versus host reaction also has important clinical implications, for it can occur in the course of certain transplantation endeavors in humans, especially transplants of immunocompetent cells.[76]

In Vitro Methods. In vitro methods for the detection, study, and categorization of transplantation antigens have replaced to a great degree in vivo methods.[3] These in vitro methods utilize sensitive immunologic methods; they can be adapted to easily obtainable human cells; they can be used to identify the transplantation antigens in fractionated cellular material; and, probably most relevant to our interest in transplantation, they permit us to tissue type a large number of potential organ donors and compare those results with the tissue type of a needy recipient. Numerous immunologic techniques have been applied to these in vitro studies of transplantation antigens, but a description of two will serve to illustrate the principles involved, their limitations, and their use in human organ transplantation.

Transplantation antigens are present on the surfaces of most cells of the body. The choice of representative cells for typing depends on several considerations. The cells must carry all the important transplantation antigens, they must be readily available, and they must lend themselves to immunologic techniques. Peripheral white blood cells meet all these criteria and are, therefore, the cells most used for tissue typing. It is actually more accurate to speak of leukocyte typing than tissue typing, but the distinction seems trivial.

It might be of interest to note that human red blood cells do not carry the major transplantation antigens. While the red cell antigens A and B are important and must be matched in donor-recipient pairs, they are not the major transplantation antigens most critical for histocompatibility. By contrast, mouse, rat, and chicken red blood cells do carry important tissue antigens and can be used for serologic studies. From a technical point of view it is unfortunate that human red blood cells cannot be used, because red blood cells lend themselves nicely to many serologic techniques.

Lymphocytotoxicity test. The most commonly used assay for the presence of a particular transplantation antigen on a white blood cell is the lymphocytotoxicity test.[5, 72] This test requires (1) antisera to specific transplantation antigens, (2) lymphocytes, and (3) an indicator system to detect whether a particular antiserum interacted with the lymphocyte.

The most crucial aspect of this test is the antisera. If specific antisera were available for all major human transplantation antigens, tissue

typing would be as straightforward as red blood cell typing. Unfortunately, this goal is years away. At present, antisera are available that detect many different cellular antigens, but most antisera react with a spectrum of antigens rather than with a single one. For example, one antiserum may contain antibodies to antigens 1, 2, and 3. It cannot, therefore distinguish a cell with antigen 1 from another with antigen 2, or yet another with antigen 3. A second antiserum may contain antibodies to antigens 2, 3, and 4. Lack of monospecificity coupled with overlapping specificities makes the interpretation of results difficult.

Antisera to human transplantation antigens are obtained from people who were either unintentionally or intentionally immunized with another person's transplantation antigens. Unintentional immunization occurs principally in two situations: pregnancy and blood transfusion. Most mothers are exposed to blood cells from the fetus at the time of delivery.[77, 78] White blood cells in such an unintentional transfusion possess transplantation antigens that the fetus inherited from the father. After several pregnancies a woman often has antibodies to a variety of transplantation antigens.[51] Essentially the same situation exists after repeated blood transfusions. Antisera obtained after such unintentional immunization have been used extensively and effectively to characterize the transplantation antigens of man.[3, 4, 17] Many such antisera must be used, however, in order to compare any two individuals, for each antiserum reacts with several antigens.[75]

The other approach for obtaining antisera to transplantation antigens is to intentionally immunize one individual with white blood cells from another.[55] This technique has proved safe and effective. It has resulted in the generation of antibodies against only the transplantation antigens of one other person's cells. The numbers of antigens used in the immunization can be even further reduced if daughter or son to father immunizations are performed. In this situation the father will make antibodies only to the antigens coded by one half of the mother's chromosomes. This technique, coupled with absorption, has produced the most specific antisera available. Antisera which react with discrete transplantation antigens are obviously extremely valuable in determining the antigenic makeup of a cell. Unfortunately, they are extremely rare.

The possibility of obtaining useful antisera to human transplantation antigens by immunizing various laboratory animals is very attractive, but unfortunately impractical at this time. For example, when rabbits are injected with human leukocytes they produce antibodies to the myriad of antigens on the leukocytes. There are so many antigens on the cells that are common to all human leukocytes that the resulting antiserum cannot distinguish the transplantation antigens. Humans, by contrast, do not produce antibodies to those antigens on another person's cells that are shared by the species. When transplantation antigens can be completely isolated from all other cellular antigens, then antibodies with great specificity can be generated in laboratory animals.[9] Until such time, human antisera must suffice.

Blood lymphocytes are the representative cells tested.[5] They are separated from other blood cells by simple physical methods. Care must

be taken, of course, to standardize their preparation in order to assure reproducible results. Lymph node, spleen, and thoracic duct lymphocytes can also be used in this test.

The indicator system to detect whether a particular antiserum reacts with a lymphocyte works in the following way: Aliquots of lymphocytes are added to a series of plastic wells; each well contains a different antiserum and a source of complement. The antisera-cell combinations are incubated for 30 minutes; trypan blue dye is then added. Cells that are attacked by antibody are killed by the interaction of antigen, antibody, and complement. Dead cells stain blue; live cells remain colorless. Wells containing dead cells are termed positive. The system is presented schematically in Figure 1.

A representative study of lymphocytes from two individuals, A and B, might appear as follows:

	Antiserum Number									
	1	2	3	4	5	6	7	8	9	10
A	+	0	0	+	+	+	0	+	0	0
B	+	+	0	0	+	+	+	0	0	+

It is apparent that lymphocytes from A and B are not killed by the same panel of antisera. This implies that A's and B's cells have different antigens on their surfaces. Interpretation beyond this point is possible but treacherous.

The problem of the antisera, mentioned briefly before, places the greatest limitation on further interpretation. If all 10 antisera were monospecific—that is, reacted with only one transplantation antigen—things would be fine. Then A and B could be said to be identical for antigen 1, disparate for 2, both lack 3, etc. For reasons stated previously, this situation does not generally obtain.

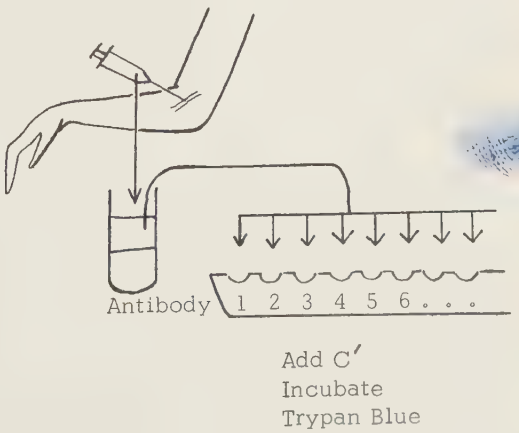


Figure 1. Schematic diagram of the indicator system used to detect whether a particular antiserum reacts with human lymphocytes.

Greater knowledge of the individual antisera is, therefore, essential for further interpretation. Whether an antiserum was generated by intentional or unintentional immunization is important to know. Intentional immunization produces antisera that react with a smaller number of antigens than do antisera obtained from humans after exposure to cells from several individuals. It is also important to know whether an antiserum had been absorbed with lymphocytes of people other than the one used for immunization. Such absorptions reduce the number of antigens with which an antiserum will react. The comparison of a particular antiserum with others also helps categorize its specificity. If a panel of 50 different lymphocyte populations is tested with an unknown antiserum plus several other antisera, it can be determined whether the unknown antiserum reacts with the same or different populations of cells than the other antiserum. Knowledge of these and other parameters improves the validity of the interpretation of tissue typing results.

It is not known today how many different human transplantation antigens exist. It is generally believed that available antisera permit delineation of 11 antigens coded by the major histocompatibility loci designed HL-A, but no doubt many others exist. If and when the categorization of human transplantation antigens is relatively complete, then the importance or strength of the various antigens will have to be determined. Clearly, if we are to improve our tissue typing capabilities and sharpen the clinical decisions that are dependent upon them, there is a pressing need to further define human transplantation antigens.

Mixed lymphocyte culture. A second in vitro test, the mixed lymphocyte culture (MLC), has proved extremely useful for both fundamental and applied study of transplantation antigens.[1, 7] This test is based on the principle that when lymphocytes are exposed to the antigens of foreign cells they are stimulated to proliferate.[6, 8, 53] The amount of proliferation occurring in a population of lymphocytes mixed with leukocytes from another person is proportional to the antigenic disparity existing between the two populations. If no lymphocyte proliferation occurs, it can be concluded that no major antigenic differences exist between the two populations of cells.

A brief description of an actual test may help to clarify the usefulness and the limitations of this approach to tissue typing. Blood is drawn from persons A and B, defibrinated, and allowed to sediment. The white blood cells from each individual are removed, washed, and dispensed into two tubes. Cells in one tube from each person are treated briefly with mitomycin to block DNA synthesis by those cells. The mitomycin-treated cells are washed to remove the extra mitomycin, and subsequently are designated Am and Bm. Four different combinations of cells are then set up for culture: A + Am, A + Bm, B + Bm, and B + Am. The cultures are incubated under appropriate conditions for 7 days. Four hours before harvesting the cultures, radioactive thymidine (^{3}H-Thy) is added to each culture. Later the nucleoprotein from each culture is extracted and the amount of radioactivity in it is counted. DNA synthesized during the last 4 hours of culture will contain a radioactive label. The amount of radioactivity is proportional to the amount

of lymphocyte proliferation in the culture; this, in turn, is proportional to the antigenic disparity between the two populations of cells in the culture tube. The procedure is schematically represented in Figure 2.

Results of such a culture expressed as counts per minute of ^{3}H may appear as follows:

		Stimulating Cells	
		Am	Bm
Responding cells	A	(100)	5000
	B	7000	(150)

This indicates that Bm stimulated A and Am stimulated B. The controls, Am + A and Bm + B, are in parentheses. These controls reflect the amount of DNA synthesis that occurred in A and B cells not stimulated by a foreign cell. Actual counts may vary considerably from one experiment to another, but triplicate cultures and other controls permit interpretation.

If A had an identical twin brother, AT, then Am + AT or A + ATm would both be at background count level. Actual family studies reveal that 25 per cent of siblings fail to stimulate one another in this MLC test, even though they are not twins.[1] This observation has great significance and will be discussed later.

The beauty of the MLC test is that it does not depend on antisera. It can reveal disparities between two persons that might well be missed

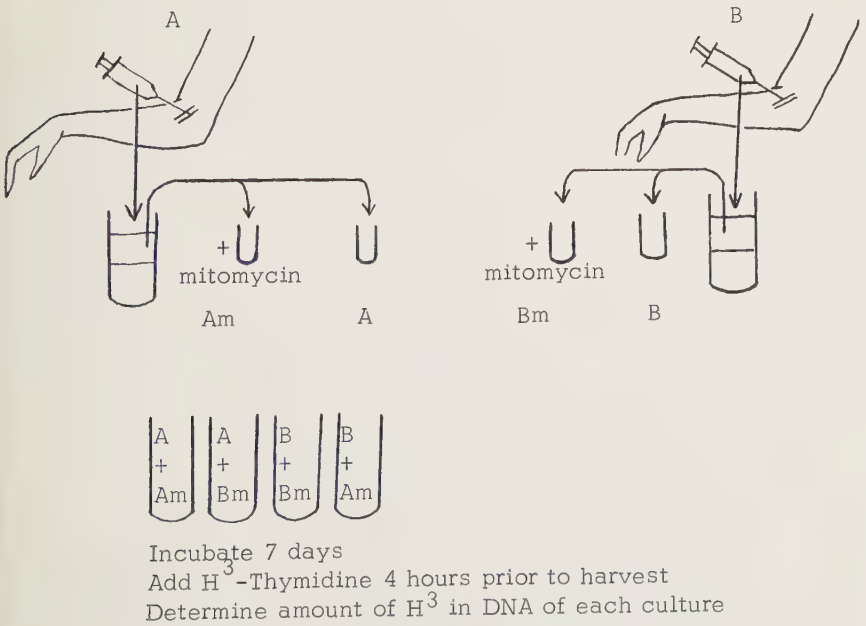


Figure 2. The mixed lymphocyte culture test for the study of transplantation antigens.

by the incomplete analysis resulting from serotyping alone. The combined used of the lymphocytotoxicity test and the MLC provides the most accurate approach to tissue typing and donor pair selection available today. The special usefulness of this combination when studying the distribution of transplantation antigens within families will become apparent in the next section.

Inheritance of Transplantation Antigens

The truly giant step in transplantation biology was the discovery that transplantation antigens, or susceptibility factors as they were called at the time, were regulated by the genetic material of an animal.[31, 35, 36] The subsequent studies defining the patterns of inheritance of the histocompatibility genes have been mainly hard work coupled with considerable ingenuity.

The genetic aspects of transplantation antigens are complex, and the complexity is compounded by the fact that large pieces of fundamental knowledge are missing. In this review of tissue immunity and tissue typing it is difficult to know how deeply to explore the subject. A full exploration would discourage all but the most enthusiastic geneticist, while a superficial one runs the risk of oversimplifying the situation and implying that solutions are almost at hand. Perhaps we can stay between these two extremes.

First, let us eliminate from the discussion the human red cell antigens A and B. These antigens are important transplantation antigens and they are, of course, coded for by inherited DNA. Nevertheless, they are not the "tissue" antigens that are generally referred to when discussing transplantation antigens. The many other antigens present on red cells do not seem to be important transplantation antigens.

The major tissue antigens that elicit immunity are coded for by genetic information located at various sites on the chromosomes. In the mouse, where breeding experiments have permitted exacting study, there is a series of genetic loci that code for transplantation antigens.[64] Thirteen loci have been described, designated H-1 to H-11 for autosomal loci and X and Y for two sex-linked loci. The H-2 locus is responsible for the major transplantation antigens—that is, those antigens that must match if tissue grafts are to be successful. The H-2 locus is on an autosomal chromosome of the ninth linkage group. Each diploid cell thus has two H-2 loci, one from each parent, and the products of both loci are expressed in each cell. There are also a number of subloci at the H-2 locus, each of which has multiple alleles. Altogether there are at least 33 antigens known to be coded by the H-2 locus. Various combinations of these antigens make possible hundreds of different cell types.

Antigens coded by other histocompatibility loci are not nearly as "strong" as those coded by H-2. Transplants made across non–H-2 differences are rejected with less vigor than those differing by H-2 loci.

The situation in man appears quite comparable to that in the mouse insofar as the evidence goes.[18, 19, 59] The major transplantation locus in man, designated HL-A, codes for a series of as yet incompletely defined antigens. The locus is generally inherited as a unit and both genes are

expressed on the surface of a cell. In other words, the genes are codominant. The in vitro assays of human transplantation antigens coupled with actual transplants, i.e., skin or kidneys, seem to confirm the similarities between H-2 and HL-A. This apparent similarity requires much more documentation, however, before it can be fully accepted.

As the HL-A locus is inherited as a unit and each locus is codominant, a family can be schematically represented as follows:

Mother		*Father*
AB	×	CD

AC BC AD BD

Only four types of sibs are possible, AC, BC, AD and BD. Twenty-five per cent of sibs inherit identical HL-A genes. Such sibs may be quite different in every other way, but they are HL-A identical. Tissue transplants between such HL-A identical sibs elicit less immunity and therefore experience less rejection than do those between sibs who differ by both genes. Parent-child transplants or transplants between sibs differing by one gene experience intermediate rejection phenomena.

The tissue typing techniques required to sort out the family in this way are available. The mixed lymphocyte culture (MLC) appears to reveal only HL-A coded antigens and thus is negative in HL-A identical sibs. By manipulating the cell concentrations used in the MLC it is possible to distinguish one-gene-different sibs from two-gene-different sibs. Parent's cells stimulate to the same degree as those of a one-gene-different sib.[1]

The lymphocytotoxicity test also can sort out the inheritance pattern of the HL-A locus within a family. This test coupled with the MLC permits an accurate assessment of transplantation antigens within a family and enables accurate matching of tissue transplants among sibs.

The human, as the mouse, also has transplantation antigens coded by genetic loci other than HL-A. These other loci are not inherited on genes linked to HL-A and thus will be disparate even though HL-A genes are identical.

The necessity remains to improve and apply serologic techniques to large familes and unrelated populations in order to definitively map the human transplantation genes, determine the number of subloci and alleles at HL-A, and develop antisera capable of fully characterizing this important locus. Until this is accomplished, efforts to "match" transplantation antigens between unrelated persons will always have to be qualified.

MECHANISM OF GRAFT REJECTION

Tissue antigens apparently immunize an animal in much the same way as do other foreign antigens. The transplantation antigens are recognized as "foreign" and trigger off the sequence of events characterized as the immune response. This is an extremely glib explanation, of course, for many of the most important steps in the "immune re-

sponse" are speculative at best. Let us just concentrate on a few special features of tissue immunity and graft rejection.

First, tissue antigens are generally thought of as fixed to the tissues; therefore, considerable curiosity has existed concerning how these antigens stimulate the immune system of a new host. Two ways have been envisioned and both probably are in fact instrumental.[45] Transplantation antigens can come off the transplanted tissues and migrate to regional lymph nodes or spleen, there to stimulate immunocompetent lymphocytes.[46] Immunocompetent lymphocytes circulating in the blood may also meet the foreign transplantation antigens antigens in the transplanted tissue.[68] This latter situation is especially likely in an immediately vascularized graft such as a kidney or heart.

The immune response to foreign transplantation antigens is manifested in two ways, cellular immunity and antibody synthesis. Cellular immunity is the phenomenon wherein lymphocytes become sensitized to a substance and directly interact with the substance when it is next emcountered. This direct interaction of a sensitized lymphocyte and antigen causes a delayed type of inflammatory response most classically seen in the tuberculin reaction. This direct interaction also seems to characterize most graft rejection reactions.[13, 26, 42] Serum antibody and complement do not appear to be involved. The mechanism whereby a lymphocyte becomes sensitized and able to interact with only a specific antigen is a mystery.

Antibody synthesis also occurs in response to foreign transplantation antigens.[27, 57] Serum antibodies can be instrumental in graft rejection although cellular immunity appears to play a more prominent role. The final definition of the mechanism of cellular immunity and the relative importance of cellular and humoral immunity in tissue rejection awaits the results of more fundamental studies of immunity.

IMMUNOLOGIC TOLERANCE

It is startling how many fundamental biologic phenomena can be discussed and even applied to human medicine without being thoroughly understood. Immunologic tolerance is one of these. It refers to specific immunologic unresponsiveness to substances that would generally be antigenic and elicit an immune response; it is induced in an animal by prior exposure to the substance.[11, 28, 40, 48] Transplantation antigens are especially important substances to consider in this regard. An animal will not develop an immune response to transplantation antigens that are present during his development. Owen first observed this phenomenon when he noted that freemartin cattle, dizygotic twins with a common blood supply, shared one another's red blood cells all through life.[50] They tolerated antigens that under other circumstances would have elicited an immune response and been rejected. Medawar later showed that newborn mice exposed to foreign transplantation antigens would not reject tissues bearing those antigens, though an adult would.[40] By various manipulations, including x-rays and immunosuppressant

drugs, it is possible to induce immunologic tolerance to transplantation antigens even in an adult animal.[23, 37, 54, 74]

Although the mechanism of tolerance remains unknown, its applicability to tissue transplantation is obvious. For example, if antigens coded by the H-2 locus of two mice can be matched, it is quite simple to develop tolerance to the other transplantation antigens. The ultimate goal in human tissue transplantation endeavors is to match the HL-A coded antigens as closely as possible and establish tolerance to the others.

BIOLOGICAL SIGNIFICANCE OF TISSUE IMMUNITY

The subtle but extremely critical mechanisms that must exist for the homeostasis of rapidly replicating cells within an animal are probably closely linked to transplantation antigens and tissue immunity. Each individual has a unique pattern of antigens on his cell surfaces plus an immune system sensitively attuned to that pattern. This combination has all the makings of an ingenious homeostatic mechanism.[15, 73] Many years ago Thomas said, "tissue immunity did not evolve to frustrate the transplant surgeon."[73] The more we learn of tissue immunity the more poignant this facetious remark becomes. Even the slightest change in cell surface antigens can be detected by the alert immune system. Those altered cells are then legitimate antigens and ultimate victims of tissue immunity.

Most cancer cells have new antigens on their surfaces and are thus foreign to an animal.[49, 66] Increasing the immune response to cancer cells in humans, by a number of techniques, occasionally has been followed by the disappearance or the diminution of a tumor.[61, 65] This approach to cancer therapy is promising.[43, 44, 60] On the other hand, diseases or medications that impair the immune response predispose to the appearance of cancer. Immunologic deficiency diseases and immunosuppressent drugs are examples.[22] Aging is accompanied by a decreased immunologic responsiveness[41, 71, 79] and by an increased occurrence of cancer.[12, 24, 52, 56, 70] Perhaps the two phenomena are related.

Infectious agents interact with cells and among other things alter their transplantation antigens. Such cells are therefore foreign and subject to immunologic attack. Perhaps this process is the basis of certain autoimmune disorders.[21, 31] Many examples could be given. It is quite possible that a full understanding of tissue immunity will one day clarify many mysteries now existing in biology and in clinical medicine.

CONCLUSIONS

Tissue immunity is a fundamental biologic process. Knowledge about it is of importance to the physician, for it is the principal cause of rejection of foreign tissue transplants and it may be critical in the pathogenesis of cancer and autoimmune disorders. Many aspects of

aging may be linked to alterations in the subtle balance between transplantation antigens and tissue immunity.

Tissue typing techniques are just beginning to have an impact on medicine. One day these techniques will no doubt markedly improve tissue transplantation endeavors. It is hoped that they will also lead to a better understanding of the fundamental changes accompanying cancer, infection, and aging.

REFERENCES

1. Albertini, R. J., and Bach, F. H.: Quantitative assay of antigenic disparity at HL-A—The major histocompatibility locus in man. J. Exp. Med., *128*:639, 1968.
2. Amos, D. B.: Transplantation antigens in mouse, rat, and man. *In* Progress in Medical Genetics, 3:106, 1964.
3. Amos, D. B.: Determination of histocompatibility. The inheritance of leukocyte antigens. Transplantation, 5:1015, 1967.
4. Amos, D. B., et al., eds.: Histocompatibility Testing 1965. Copenhagen, Munksgaard, 1965.
5. Amos, D. B., Bashir, H., Boyle, W., MacQueen, M., and Tüilkainen, A.: A simple micro cytotoxicity test. Transplantation, 7:220, 1969.
6. Bach, F., and Hirschhorn, K.: Lymphocyte interaction: A potential in vitro histocompatibility test. Science, *143*:815, 1964.
7. Bach, F., and Voynow, N. K.: One-way stimulation in mixed leukocyte cultures. Science, *153*:546, 1966.
8. Bain, B., and Lowenstein, L.: Genetic studies on the mixed leukocyte reaction. Science, *145*:1315, 1964.
9. Batchelor, J. R., and Sanderson, A. R.: HL-A substances: Properties and immunization of rabbits. Transplantation Proceedings, *1*:489, 1969.
10. Billingham, R. E., and Silvers, W. K.: Sensitivity to homografts of normal tissues and cells. Ann. Rev. Microbiol., *17*:531, 1963.
11. Billingham, R. E., Brent, L., and Medawar, P. B.: Actively acquired tolerance of foreign cells. Nature, *172*:603, 1953.
12. Blumenthal, H. T., and Rogers, J. B.: Spontaneous and induced tumors in the guinea pig, with special reference to the factor of age. Progr. Exp. Tumor Res., *9*:261, 1967.
13. Brent, L., and Medawar, P. B.: Quantitative studies on tissue transplantation immunity. V. The role of anti-serum in enhancement and desensitization. Proc. Roy. Soc. (London), *155*:392, 1962.
14. Brent, L., and Medawar, P. B.: Tissue transplantation: A new approach to the "typing" problem. Brit. Med. J., *2*:269, 1963.
15. Burnet, F. M.: Immunological recognition of self. Science, *133*:307, 1961.
16. Converse, J. M., and Casson, P. R.: The historical background of transplantation. *In* Rapaport, F. T., and Dausset, J., eds.: Human Transplantation. New York, Grune and Stratton, Inc., 1968, p. 3.
17. Curtoni, E. S., Mattiuz, P. L., and Tosi, R. M., eds.: Histocompatibility Testing 1967. Baltimore, Williams and Wilkins, 1967.
18. Dausset, J., and Rapaport, F. T.: The Hu-1 system of human histocompatibility. *In* Rapaport, F. T., and Dausset, J., eds.: Human Transplantation. New York, Grune and Stratton, 1968, p. 369.
19. Dausset, J., Ivanyi, P., and Ivanyi, D.: Tissue alloantigens in humans. Identification of a complex system (Hu-1). *In* Amos, D. B., et al., eds.: Histocompatibility Testing 1965. Copenhagen, Munksgaard, 1965, p. 51.
20. Davies, D. A. L., Manstone, A. J., Viza, D. C., Colombani, J., and Dausset, J.: Human transplantation antigens: The HL-A (HU-1) system and its homology with the mouse H-2 system. Transplantation, 6:571, 1968.
21. Decker, J. L., ed.: The Relationship of Mycoplasma to Rheumatoid Arthritis and Related Diseases. Public Health Service Publication No. 1523. Washington, U.S. Department of Health, Education and Welfare, 1966.
22. Dent, P. B., Peterson, R. D. A., and Good, R. A.: The relationship between immunologic function and oncogenesis. *In* Good, R. A., ed.: Immunologic Deficiency Diseases in Man. New York, The National Foundation, 1968, p. 443.
23. Dixon, F. J., and Maurer, P. H.: Immunologic unresponsiveness induced by protein antigens. J. Exp. Med., *101*:245, 1955.
24. Failla, G.: The aging process and cancerogenesis. Bull. N.Y. Acad. Sci., *71*:1124, 1958.
25. Goldsmith, K. L. G.: Donor selection and compatibility typing. Brit. Med. Bull., *21*:162, 1965.

26. Govaerts, A.: Cellular antibodies in kidney homotransplantation. J. Immunol., *85*:516, 1960.
27. Hager, E. B., DuPuy, M. P., and Wallach, D. F. H.: Immunologic suicide and studies on the role of antibody and complement in canine kidney homograft rejection. Ann. N.Y. Acad. Sci., *120*:447, 1964.
28. Hašak, M., Lengerová, A., and Hraba, T.: Transplantation immunity and tolerance. *In* Taliaferro, W. H., and Humphrey, J. H., eds.: Advances in Immunology, *1*:1, 1961.
29. Hellström, K. E., and Hellström, I.: Allogeneic inhibition of transplanted tumor cells. Progr. Exp. Tumor Res., *9*:40, 1967.
30. Hellström, K. E., and Möller, G.: Immunological and immunogenetic aspects of tumor transplantation. Progr. Allergy, *9*:158, 1965.
31. Jensen, C. O.: Transplantation of mammary gland carcinoma in mice. Zentralbl. Bakt., *34*:28, 1903.
32. Kahan, B. D., and Reisfeld, R.: Electrophoretic purification of a water-soluble guinea pig transplantation antigen. Proc. Nat. Acad. Sci., *58*:1430, 1967.
33. Kahan, B. D., and Reisfeld, R. A.: Advances in the chemistry of transplantation antigens. Transplantation Proceedings, *1*:483, 1969.
34. Lawrence, H. S.: Homograft sensitivity. Physiol. Rev., *39*:811, 1959.
35. Little, C. C., cited in Snell, G. D., ed.: Biology of the Laboratory Mouse. Philadelphia, Blakiston, 1941, p. 279.
36. Loeb, L., cited in The Biological Basis of Individuality. Springfield, Illinois, Charles C Thomas, 1945.
37. Mariani, T., Martinez, C., and Good, R. A.: Studies of sex-linked histo-incompatibility in inbred mice: Role of age and production of immunological tolerance. Inter. Arch. Allergy and Appl. Immunol., *16*:216, 1960.
38. Medawar, P. B.: Immunity to homologous grafted skin. III. The fate of skin homografts transplanted to the brain, to subcutaneous tissue, and the anterior chamber of the eye. Brit. J. Exp. Path., *29*:58, 1948.
39. Medawar, P. B.: Zoologic laws of transplantation. *In* Peer, L. A., ed.: Transplantation of Tissues. Baltimore, Williams and Wilkins, 1959, Vol. 2, p. 41.
40. Medawar, P. B.: Immunological tolerance (Nobel Lecture). Nature, *189*:14, 1961.
41. Metcalf, D., Moulds, R., and Pike, B.: Influence of the spleen and thymus on immune responses in aging mice. Clin. Exp. Immunol., *2*:109, 1967.
42. Mitchison, N. A.: Passive transfer of transplantation immunity. Nature, *171*:267, 1953.
43. Moore, G. E., and McLimans, W. F.: The life span of the cultured normal cell: Concepts derived from studies of human lymphocytes. J. Theor. Biol., *20*:217, 1968.
44. Nadler, S. H., and Moore, G. E.: Immunotherapy of malignant melenoma. Geriatrics, *23*:150, 1968.
45. Najarian, J. S., and Foker, J. E.: Mechanisms of kidney allograft rejection. Transplantation Proceedings, *1*:184, 1969.
46. Nathan, P.: Antigen release from transplanted dog kidney. Ann. N.Y. Acad. Sci., *120*:485, 1964.
47. Nathenson, S. G., and Davies, D. A. L.: Solubilization and partial purification of mouse histocompatibility antigens from a membranous lipoprotein fraction. Proc. Nat. Acad. Sci., *56*:476, 1966.
48. Nossal, G. J. V.: Immunologic tolerance. *In* Rapaport, F. T., and Dausset, J., eds.: Human Transplantation. New York, Grune and Stratton, 1968, p. 643.
49. Old, L. K., and Boyse, E. A.: Immunology of experimental tumors. Ann. Rev. Med., *15*:167, 1964.
50. Owen, R. D.: Immunogenetics of vascular anastomosis between bovine twins. Science, *102*:400, 1945.
51. Payne, R.: Development and persistence of leukoagglutinins in parous women. Blood, *19*:411, 1962.
52. Sacher, G. A.: On the statistical nature of mortality with especial reference to chronic radiation. Radiology, *67*:250, 1956.
53. Schrek, R., and Donnelly, W. J.: Differences between lymphocytes of leukemic and nonleukemic patients with respect to morphologic features, motility, and sensitivity to guinea pig serum. Blood, *18*:561, 1961.
54. Schwartz, R.: Drug induced immunological tolerance. J. Clin. Invest., *38*:1041, 1959.
55. Shanbrom, E., Feingold, E., Shepherd, L., and Walford, R. L.: Antisera for tissue typing: The production of cytotoxic and agglutinating antibodies following intradermal leukocyte injections in man. Blood, *32*:402, 1968.
56. Simms, H. S., and Berg, B. N.: Longevity and the onset of lesions in male rats. J. Geront., *12*:244, 1957.
57. Simonsen, M.: Biological incompatibility in kidney transplantation in dogs. II. Serological investigations. Acta Path. et Microbiol. Scandinav., *32*:36, 1953.
58. Simonsen, M.: Recent experiments on the graft-vs-host reaction in the chick embryo. Brit. Med. Bull., *21*:129, 1965.
59. Simonsen, M.: Strong transplantation antigens in man. Lancet, *1*:415, 1965.

60. Skurkovich, S. V., Makhonova, L. A., Reznichenko, F. M., and Chervonskiy, G. I.: Treatment of children with acute leukemia by passive cyclic immunization with autoplasma and autoleukocytes operated during the remission period. Blood, *33*:186, 1969.
61. Smith, R. T.: Tumor-specific immune mechanisms. New Eng. J. Med., 278:1207, 1268, 1326, 1968
62. Snell, G. D.: Methods for the study of histocompatibility genes. J. Genet., *49*:87, 1948.
63. Snell, G. D.: The genetics of transplantation. J. Nat. Canc. Inst., *14*:691, 1953.
64. Snell, G. D., and Stimpfling, J. H.: Genetics of tissue transplantation. *In* Green, E. L., ed.: Biology of the Laboratory Mouse. 1966, p. 457.
65. Southam, C. M.: Applications of immunology to clinical cancer—past attempts and future possibilities. Cancer Res., *21*:1302, 1961.
66. Southam, C. M.: Evidence for cancer-specific antigens in man. Prog. Exp. Tumor Res., *9*:1, 1967.
67. Streilein, J. W., Hildreth, E. A., Ramseier, H., and Kornblum, J.: The irradiated hamster test. A new method of donor selection for homotransplantation in man. Ann. Intern. Med., *65*:511, 1966.
68. Strober, S., and Gowans, J. L.: The role of lymphocytes in the sensitization of rats to renal homografts. J. Exp. Med., *122*:347, 1965.
69. Summerell, J. M., and Davies, D. A. L.: Further characterization of soluble mouse and human transplantation antigens. Transplantation Proceedings, *1*:479, 1969.
70. Szilard, L.: On the nature of the aging process. Proc. Nat. Acad. Sci., *45*:30, 1959.
71. Teller, M. N., Stohr, G., Curlett, W., Kubisek, M. L., and Curtis, H.: Aging and cancerigenesis. I. Immunity to tumor and skin grafts. J. Nat. Cancer Inst., *33*:649, 1964.
72. Terasaki, P. I., and McClelland, J. D.: Microdroplet assay of human serum cytotoxins. Nature, *204*:998, 1964.
73. Thomas, L., *In* Lawrence, H. S., ed.: Cellular and Humoral Aspects of the Hypersensitivity States. New York, Hoeber Medical Division, Harper & Row, 1959, p. 529.
74. Trentin, J. J.: Induced tolerance and "homologous disease" in x-irradiated mice protected with homologous bone marrow. Proc. Soc. Exp. Biol. N.Y., *96*:139, 1957.
75. Troup, G. M., Ramsey, N., Flynn, R., and Walford, R. L.: The production of lymphocytotoxic antisera of narrowed specificity by selective adsorption and immunization. Vox Sang., *11*:315, 1966.
76. Uphoff, D. E.: Immunological considerations of bone marrow transplantation. Transplantation Proceedings, *1*:39, 1969.
77. Weisert, O., and Marstrander, J.: Severe anemia in newborn caused by feto-maternal "transfusion." Acta Paediat., *49*:426, 1960.
78. Wiener, A. S.: Diagnosis and treatment of anemia of newborn caused by occult placental hemorrhage. Amer. J. Obstet. Gynec., 56:717, 1948.
79. Wigzell, H., and Stjernsward, J.: Age-dependent rise and fall of immunological reactivity in the CBA mouse. J. Nat. Cancer Inst., 37:513, 1966.
80. Wolf-Jürgensen, P., and Schwartz, M.: Normal-lymphocyte transfer in man. Basophil leukocytes in delayed skin reaction. Lancet, 2:388, 1964.

La Rabida Children's Hospital and Research Center
East 65th Street at Lake Michigan
Chicago, Illinois 60649

Allergic Reactions to Antibiotics

JOHN E. KASIK, M.D., Ph.D.*

JOHN S. THOMPSON, M.D.**

The widespread use of antibiotics in clinical medicine, in veterinary therapy, and in agriculture has exposed a large number of individuals to a diverse array of potentially sensitizing agents. As a result, an allergic reaction to antibiotics has become a problem which must be faced by anyone who employs these substances. This adverse reaction must be considered as a risk to the patient before treatment and weighed against the potential therapeutic benefits. Frequently, a history of a previous hypersensitivity reaction to an antibiotic will be elicited from a patient and preclude the use of an agent with compelling therapeutic advantages. Even when the history is dubious and the need for the antibiotic is great, the physician finds himself in a difficult situation where the only satisfactory way to prove or disprove hypersensitivity is to use the drug and await a reaction.

Reactions to antibiotics can also be a serious medicolegal problem. The injudicious use of one of them may be almost as risky to the physician as to his patient. Using antibiotics without an indication, failure to identify a reaction when it occurs, and failure to treat reactions promptly and properly can lead to serious effects in the courts. The physician who uses these substances must be aware of his indications, diagnose reactions promptly and treat them properly to avoid this other, secondary, adverse reaction.[56]

This review is an attempt to clarify the problem of adverse reactions to antibiotics involving immunological mechanisms, describing the forms the disorder may take, indicating the risks involved, discussing the mechanisms responsible, where known, and outlining the methods of prevention and treatment for allergic reactions to the more commonly used antibiotics.

*Associate Professor of Medicine and Pharmacology, University of Chicago, Chicago, Illinois
**Professor of Medicine, University of Iowa, Iowa City, Iowa

Prepared with the support of grant No. AI-08517-01 from the U.S. Public Health Service, and a grant from the Illinois Tuberculosis Association.

Manifestations and Incidence

All the clinically available antibiotics have been reported to have produced hypersensitivity reactions, but the incidence and severity of the problem varies widely according to the antibiotic. In general the reactions fall into two broad classifications, immediate and delayed, with the former potentially more serious because of the association with anaphylactoid reactions.[40] Because of the earlier recognition of an allergic reaction to biologicals containing horse serum, the term "serum sickness type" has been applied to some forms of the delayed type of reaction.[20] The delayed type usually occurs some days after initiation of therapy, but it is well recognized that the division between delayed and immediate reactions is not clear and that both can result in similar manifestations. Both are mediated by humoral antibodies.

The commonest organ system involved in the delayed form of reactions has been the skin,[43, 60] with reactions occurring alone or with involvement of other organ systems. These reactions run the gamut from minor erythema at the injection site to a fulminant bullous reaction. Among the skin manifestations of allergic reactions that have been reported with antibiotic therapy are erythema, erythematous rashes, urticaria, pruritus, angioedema, lichen planus-like eruptions, bulla formation, and exfoliative dermatitis. The rashes can be morbilliform or scarlatiniform and petechial rashes without thrombocytopenia have been reported—especially with penicillin, streptomycin, and the tetracyclines.[13] Vasculitis, noted most frequently with penicillin, may lead to hemorrhage or ulceration.

An unusual adverse effect of antibiotics manifest in the skin which may be an allergic reaction is the so-called photosensitivity reaction. It has been observed with the tetracyclines, especially demethylchlortetracycline, and the griseofulvins.[2, 29] These reactions appear in susceptible individuals after drug ingestion and exposure of the skin to a suitable wavelength of light, usually direct sunlight. The reaction occurs in the exposed areas of the body and occurs as an exaggerated sunburn-like response, with intense erythema followed by vesicles, edema, and the formation of bullae in the skin after minimal exposure to sunlight. Nail discoloration has also been noted, sometimes with loss of the nail at the nail bed.

Contact dermatitis from antibiotics, either applied as a medication or as the result of skin contact in medical personnel, is a well recognized allergic problem. In the past it has been especially prominent with neomycin preparations used in topical therapy and as the result of contamination of the hands with streptomycin in nurses and physicians.[8] While this problem has been widely discussed in the past and the dangers of using common and potentially life-saving antibiotics as topical medications have been stressed and the practice condemned, gentamicin was recently introduced into therapy as a topical medication despite the early indications that this antibiotic could be a substantial addition to systemic therapy of serious gram-negative infections.

Myalgia and joint pains, sometimes with swelling, are frequent in allergic drug reactions.[34] In some instances findings compatible with a

collagen disease-like picture have been observed. A clinical picture resembling disseminated lupus erythematosus has been noted in severe delayed reactions to antibiotics, particularly to benzylpenicillin and, in one or two instances, to tetracycline, streptomycin, and the antifungal agent griseofulvin.[31] In this situation, such well recognized lupus-associated phenomena as urine sediment findings, biologically false positive serologies, and positive LE preparations have been noted. This reversible syndrome has been called the pharmacologic systemic lupus syndrome by Symmers.[54] Biological false positive serological tests for syphilis have also been noted after recovery from penicillin reactions without other signs of LE.[9]

Fever, as an aspect of hypersensitivity reaction to drugs, can be a perplexing clinical problem[11, 34] and a pitfall for the physician in a clinical situation in which the patient is being treated for a bacterial infection. Drug fever, as a result of hypersensitivity to antibiotics, has been reported with a variety of antibiotics, including penicillin and streptomycin, and may present in a variety of ways. The reaction may appear as a part of a delayed or immediate allergic reaction and may or may not be associated with other manifestations. The fever can be sustained or remittent and the problem is sometimes associated with eosinophilia, but sometimes this clue to the diagnosis may be absent. To further complicate the clinical picture, the fever may present together with leukocytosis and lymphadenopathy. Less of a diagnostic problem is the more common situation in which the joints, skin, or other organ systems are also involved. (For review of the problem of drug fever, see Cluff and Johnson.[11])

Hematopoietic manifestations of allergy to an antibiotic have been reported. Eosinophilia is the commonest, and, as mentioned previously, leukocytosis can occur. Less common are the hemolytic anemias that have been reported with the penicillins and cephalothin and very rarely with other antibiotics such as the tetracyclines.[36, 48, 59, 62] In some instances, hemolytic problems appeared without systemic manifestations.[59] It has been reported that red cells from a patient with penicillin-induced hemolysis have been shown to cross react with ampicillin and methicillin[62] and cephalothin.[36] Other hematopoietic disturbances occur rarely. Thrombocytopenia and leukopenia as a part of penicillin allergy have been described,[48] and there has been speculation that aplastic anemia associated with chloramphenicol may be the result of a hyper-immune reaction involving this drug.

The immediate allergic reactions to antibiotics are a constellation of symptoms, in many instances similar to those described above. The reaction may occur in all degrees of severity and may be expressed in one or more organ systems, with the skin the commonest and most evident organ involved. The most violent and most serious form of an immediate allergic reaction to an antibiotic is the anaphylactoid reaction. This reaction can begin minutes or hours after administration of a substance and frequently is heralded by a feeling of uneasiness, weakness, and sweating, followed by pruritus and wheal formation, either generalized or local. Nasal itching and swelling of the face, hands, and other

soft tissues occur rapidly. Ominous signs are tachycardia, wheezing, laryngeal stridor, dyspnea, and cyanosis. Central nervous system signs may develop, with inappropriate behavior, convulsions, and coma. Shock may develop, and the signs associated with shock are usually present. Nausea, vomiting, and diarrhea may also ensue, especially after oral administration of the offending antibiotic. This violent reaction has been reported with all or nearly all the antibiotics,[1, 34, 48] but benzylpenicillin and its salts, given by injection, have been by far the substances most often responsible for the difficulty.

In Welch's classical study of reactions to antibiotics, 809 cases had an over-all mortality of about 10 per cent.[60] Of the 809 reactions reported, 793 were due to penicillin (in 1957) and 611 of those to injected penicillin. Oral preparations of penicillin were implicated in 35 cases, but it is of interest to note that no deaths were reported in this group. Subsequently, deaths have been noted with the oral penicillins by others, but the numbers remain low. Scattered reports of violent reactions with topical applications of penicillin have also appeared.[34]

The incidence of immediate reactions to penicillin has been reported to be 1.4 per 1000 patients with no prior history of allergy to penicillin.[35] This incidence, if taken with the mortality data for anaphylactic reactions, would indicate roughly 1 death in 10,000 patients treated, a figure which is certainly too high. About 300 deaths per year have been attributed to allergic reactions to penicillin in the United States.[34]

Less severe reactions to penicillin have been reported to occur in 1 to 10 per cent of patients given the drug.[46, 47, 58] Such factors as history of allergy to other substances, renal disease, abnormal liver function tests, and the male sex have been stated to increase the incidence of allergic reactions to penicillin. As would be expected, patients with a history of prior reactions to penicillin were subject to further allergic difficulties (29 per cent).

The oral penicillins and the semisynthetic penicillins such as ampicillin, methicillin, and oxacillin have been reported to produce essentially the same gamut of allergic reactions as benzylpenicillin, but the incidence appears to be lower than with the older drug.[34, 48]

Contrary to the original impression that cross allergy between penicillin and the cephalosporins was not a significant problem,[25, 26] evidence now exists that patients who have a history of penicillin reaction have a higher risk of allergic reactions to cephalothin than those who do not.[10, 18, 55] Among 54 patients who were receiving cephalothin, 7 had allergic reactions to the drug, 2 of which were anaphylactoid in nature. Five of the 7 reactors had histories of previous allergic reactions to penicillin. Nine others who also had similar histories received cephalothin and had no reaction, even one who was given the antibiotic a day after the penicillin reaction.[55]

This information must be tempered by the fact that it is well recognized that patients with clear histories of penicillin reaction frequently fail to have a discernible reaction to the antibiotic on subsequent administration. It has been reported that in patients with clear histories of penicillin allergy and who also had positive penicilloyl polylysine skin

tests who were given penicillin, roughly three fourths failed to have an evident reaction to the drug.[7]

There is as yet insufficient evidence available to assess the incidence of allergic reactions to cephaloridine and to the new oral cephalosporin, cephalexin. In one report of 149 patients receiving cephaloridine, seven had cutaneous allergic eruptions (4 per cent).[6]

The aminoglycosides are an important group of antibiotics which includes streptomycin, neomycin, kanamycin, and gentamicin. These antibiotics all have similar chemical structures and side effects. Allergic reactions to streptomycin have been reported both after injection and by contact.[34, 48] In long-term therapy, an incidence of hypersensitivity of 10 per cent has been observed.[5] Cutaneous eruptions and drug fever seem to be the most common manifestation.

Outside of scattered reports of allergic problems due to the other members of the aminoglycosides, no extensive incidence data is available for these other, newer, and less widely used antibiotics. Because of the structural similarities of the compounds, there is a theoretical possibility of cross sensitivity from one member of the group to some or all members in allergic patients. This phenomenon of apparent cross sensitivity has been reported in patients who had developed cutaneous allergic reactions as a result of the widespread practice of the topical use of ointments containing neomycin.[38, 39] Cross sensitivity was frequently noted in these individuals to kanamycin and gentamicin, but not streptomycin, applied as a patch test. These observations are of some clinical importance, since gentamicin might be considered as an alternate topical antibiotic in a neomycin-sensitive patient; on the basis of this evidence, gentamicin should not be so used.

Chloramphenicol and the tetracycline antibiotics rarely cause hypersensitivity reactions, though, as with all antimicrobials, scattered reports of reactions of all types, including anaphylaxis, exist.[48]

The macrolide antibiotics are a series of compounds of which only erythromycin and oleandomycin are clinically important. These compounds are infrequent causes of allergic reactions, which, if they occur, are mainly composed of skin rashes.

An unusual (for an antibiotic) hypersensitivity reaction has been observed with certain salts of erythromycin and triacetyloleandomycin. Erythromycin estolate, the lauryl sulfate salt of the propionic acid ester of erythromycin, is unique among the salts of erythromycin in that it has been implicated in a number of instances of cholestatic jaundice,[22, 25] perhaps analogous to the type of hepatotoxicity produced by the phenothiazines. A similar picture has been noted with triacetyloleandomycin.[48, 57] This phenomenon has been reported to occur only in about 0.5 per cent of patients receiving erythromycin estolate and is a problem usually after prolonged therapy (> 12 days).[17]

Macrolides-induced jaundice usually presents with anorexia, nausea, vomiting, abdominal pain, fever, leukocytosis, and an enlarged liver. Alkaline phosphatase, direct acting bilirubin, and transaminases are usually elevated and not infrequently peripheral eosinophilia is present. Liver biopsy demonstrates cholestasis and periportal cellular infiltrates

of mononuclear cells and eosinophils. Recovery occurs spontaneously when the drugs are discontinued, but may be slow (2 to 5 weeks).

Two infrequently used antibiotics deserve special mention because of the high incidence of skin rash, believed to be allergic in nature, associated with their use. Ristocetin has been noted to produce this problem in 4 to 20 per cent of treated individuals and a drug fever in about 5 per cent. Allergic problems with vancomycin are similar, but somewhat less frequent.[34]

Mechanisms of Hypersensitivity Reactions

In general, the specific antigen of allergic reactions to drugs may be any one of a number of substances: (1) the drug itself; (2) a drug-protein complex with the drug acting as a hapten; (3) a degradation product of the drug; or (4) a complex of drug degradation product and a protein.[45] Since antibiotics are in general low molecular weight materials, it has been generally considered that the antibiotic itself acting as an antigen is unlikely and that the latter three possibilities are the most likely mechanisms responsible for antibiotic reactions involving antigens and antibodies. The degradation products of antibiotics can also be low molecular weight substances, but the polymerization of these reactive substances may in theory form larger molecules which have antigenic capacity. In addition to these possibilities, it should also be mentioned that antibiotics are the product of fermentation and are isolated from complex media containing microorganisms and their metabolic products. The isolation of the antibiotic may permit the carryover of minute amounts of these potentially antigenic substances as contaminants that can cause reactions.

The most extensively studied allergic reaction to an antibiotic has been the hypersensitivity to penicillin, and the literature in this area of interest is voluminous (see reviews by DeWeck and Schneider[14] and Levin[32]). Originally it was believed that the mechanism of penicillin allergy was that penicillin acted as a hapten, binding to a protein or other macromolecule to form a macromolecular complex with antigenic potential. This opinion was based on observations discussed above indicating that a low molecular weight substance such as penicillin lacked innate antigenic activity and that binding by covalent bonds to larger molecules was usually essential for immune reactivity.[19]

This opinion of the source of the penicillin antigen has been challenged. Benzylpenicillin was found to be rather unreactive with proteins, at least with a stable type of binding and, as a consequence, interest has been directed toward other compounds which are derived from penicillin either by degradation in solution or by being carried over during the manufacturing process.[14]

Current research interest in penicillin allergy centers around a number of compounds which are contaminants of penicillin. These contaminants contain amino acid residues capable of producing immunological responses in animals and in man. They are derivatives of penicillin which arise by degradation of penicillin and then act as haptens, reacting with available proteins or other macromolecules to

form immunologically active materials in the classical way. Evidence indicates that these substances may arise during production or after.

Recently Stewart, Weston, Batchelor, and others[4, 51, 52] have demonstrated potent immunogenic macromolecular substances in commercial benzylpenicillin, 6-APA (6-aminopenicillic acid), ampicillin, cephalothin, and cephaloridine. These substances were found to be nondialyzable proteinaceous materials devoid of antibacterial activity and which produced a variety of amino acids upon hydrolysis.[53] Analysis of the amino acids produced indicated that they were not polymers of penicillin or its degradation products, but must be materials derived from the fermentation process, bound to penicillin derivatives, and carried over with penicillin during fractionation and isolation.

In the case of semisynthetic penicillins, another possible source of these substances was suggested.[51] All of these penicillins were produced from 6-APA, which is the product of the enzymatic cleavage of benzylpenicillin, using coliform bacteria as an enzyme source. It has been postulated that protein from these bacteria may also be linked to derivatives of 6-APA and persist as a contaminant in the subsequently manufactured semisynthetic penicillins such as methicillin or oxacillin.

The allergenic material in commercial penicillins has been shown to be highly reactive when skin tested in penicillin-sensitive individuals and has been used to produce a variety of immunological responses in animals. Further, it has been suggested that purified penicillins, freed of these contaminants, would be less likely to produce an allergic reaction of the immediate type[30, 53] and may be less likely to produce sensitization. The latter is an attractive idea, since hypoallergenic penicillins would be a useful answer to the clinical problem of penicillin allergy, especially if they reduced the incidence of immediate reactions. Unfortunately, animal experiments have shown that purified benzylpenicillin, treated to remove immunogenic macromolecular residues, is as immunogenic as the commercial benzylpenicillin from which it is prepared,[53] and this evidence implies that purified penicillins would still be capable of acting as inducers of allergy. In addition, it has been shown that the purified penicillins are still capable of producing serum sickness type of allergic responses.[61]

The reasons for these observations may be explained by the fact that it has been well established that penicillin and many of its derivatives undergo spontaneous degradation into reactive compounds, which, because of their reactivity, act as haptens with macromolecules. These substances have been extensively studied, and it has been established that many of them can react with themselves or with amino acids such as lysine or with the sulfhydryl group of cysteine to form immunogenic materials.[14, 32] In addition, under certain conditions penicillin itself can act as an inducer in antigen formation by reacting with amino acids and then can undergo further degradation.[3]

As a consequence, the number of reactive compounds that can be formed from penicillin and that can act as inducers or elicitors of penicillin allergy is substantial. Because of the relatively large amounts formed, a derivative of penicillic acid, the benzylpenicilloyl determinant

(BPO), has been called the major determinant and is considered by some to be the antigen responsible for the serum sickness type of penicillin reactions.[32] Others, the minor determinants, along with the macromolecular contaminant described previously, may be responsible for the more violent immediate reactions.[4]

These concepts, if accepted, allow a number of interesting speculations. It becomes obvious that penicillin allergy can have a large number of compounds as inducers and elicitors, and this may explain the various forms the disorder may take and the frequent failure to elicit a reaction in a known hypersensitive individual. If a particular lot of penicillin were low in antigen, or at least to the antigen to which the patient is sensitive, a reaction might not occur in that patient.

It also has been stated[32] that the relative amounts of various antigens present during the binding of antigen to antibody can be important.[14] Some of the determinants described above can bind to antibody, but do not produce an allergic reaction. By simply occupying the binding site, these compounds can prevent the binding of other antigens which are capable of producing a reaction from binding. Thus, if a significant amount of an antigen with blocking capacity were present, a reaction would not occur. Several compounds have been shown to produce this effect, including benzylpenicillin itself.

The mechanisms of allergic reactions to antibiotics other than the penicillins and the cephalosporins are obscure. Presumably they act as haptens or have metabolic products that act as haptens.[32] The mechanism of macrolide-induced hepatotoxicity has been the subject of some speculation.[22] It has been proposed that both erythromycin estolate and triacetyloleandomycin have structural similarities in the desmosomene side chain. In both cases a reactive center allows them to act as haptens. This, if correct, could explain the peculiar selectivity of these two macrolides to induce allergy, but leave open the question of why the liver is selected as a target organ.

The mechanism of drug-induced photosensitivity reactions is beyond the scope of this paper, and the reader is referred to the review by Kershbaum and Beerman.[26]

PREVENTION OF ANTIBIOTIC ALLERGIC REACTIONS

Allergic antibiotic reactions, especially reactions to penicillin, will always remain a problem. Clinicians have seen several promising methods of avoiding allergic reactions with penicillin arrive, only to fail to withstand widespread use. Nevertheless, there are several simple, currently available procedures which would prevent many allergic reactions, and some others have promise for the future.

The most elementary method of reducing the incidence of reactions to a minimum without significant effort would be to prescribe antibiotics only when clearly indicated. Anyone who is familiar with clinical medicine knows that antibiotics are frequently used inappropriately, unnecessarily, for too long a period of time, and are given parenterally

when oral preparations would be equally satisfactory. In the case of penicillin, this would seem obvious. In general, the axiom of clinical medicine would apply: potentially hazardous treatment is warranted only in a potentially serious illness.

Equally elementary is the simple precaution of carefully questioning every patient about previous allergic reactions before prescribing or injecting antibiotics. This means not only asking if the patient is allergic to any drugs, but asking specifically if he has had the drug or similar drugs before and then question him directly about pruritus, hives, swelling, or fever following the administration of the material. These latter points may seem unnecessary, but it has been reported that in 30 cases of fatal anaphylaxis after parenteral penicillin, seven patients had a clear history of a previous penicillin reaction, but no inquiry was made by the physician before injecting the drug. In three additional cases the physician claimed to have elicited a negative history of penicillin allergy from the patient, but the reverse was alleged by the deceased's family. In the same group of 30, retrospective consideration of the indications for use of penicillin led to the conclusion that only 12 out of the 30 had a clear indication for receiving the drug.[44]

It is obvious that the number of allergic reactions could be minimized by these simple precautions, but they would not be eliminated. At present, they are essentially all that can be conveniently done in the average clinical situation, especially for most of the nonpenicillin drugs. In the case of penicillin, there has been a continuing interest in various skin testing procedures with a hope that a relatively simple screening method would eliminate at least the anaphylactoid reaction. Originally, skin testing was performed using benzylpenicillin as a test substance, either by intradermal injection, as a scratch test, or by dropping the solution into the conjunctiva.[12, 49] Concentrations of 300,000 to 30,000 units per ml. or less were used, and the appearance of a wheal after 15 minutes was considered a positive test.

This method of testing has been praised[33] and condemned.[35, 58] The condemnation has been based on the incidence of false negatives (1.4 per cent) and the occasional severe reactions that occur as a result of the skin test itself.[37, 49] It has been pointed out that the incidence of reactions, after the elimination of prior reactions to penicillin by history, is essentially the same.[35] In addition, it is known that not all of those with skin test reactions to benzylpenicillin will have a reaction to the antibiotic if it is used in therapeutic doses, and as a result these people would be denied therapy with a valuable drug.

Based on current knowledge of the antigens involved in penicillin hypersensitivity, some of the problems involved in skin testing with benzylpenicillin may be the result of differences in the amounts or kinds of antigen contained in different lots of penicillin. If there were substantial differences in the antigenic content of various lots of penicillin, variation in the results of skin tests could occur.

Several methods have been devised to test with a standard antigen. Penicilloyl polylysine (PPL) is a polyfunctional hapten that has been used extensively.[35] This substance has been shown to cause skin

reactivity in 35 to 75 per cent of persons with a history of a prior penicillin reaction.[58] However, the incidence of reactions to penicillin in patients with a positive PPL skin test is not high, and while the incidence of false positive tests may not be important in less serious illnesses, it may prevent the use of penicillin in the occasional situation in which the drug is badly needed.

One theoretical objection to the use of PPL as a skin test antigen can be raised. It has been stated that the PPL-like material BPO, the major determinant, is only rarely the antigen responsible for severe immediate reactions in man and has been implicated mainly in the delayed type of reaction.[41] As a result of this observation, it may be that PPL would not reveal a significant percentage of potential immediate reactors to penicillin. In keeping with this theory that the minor determinants are responsible for anaphylaxis with penicillin, minor determinant material (MDM) has been used a a skin test antigen[32] and is reported to be a helpful addition to a screening program for potential penicillin reactors.

There are other methods which have been used to detect penicillin-sensitive individuals, such as the basophil degeneration test, but none of them are useful at present in the clinical situation.[35]

TREATMENT OF ALLERGIC REACTIONS TO ANTIBIOTICS

The treatment of allergic reactions to most antibiotics, indeed most drugs, is similar and can be generalized. In mild reactions, nothing more complicated than simple withdrawal of the drug will be sufficient. The patient should be told what has happened and warned about using the drug in the future. If necessary, he should be given a card with the name of the drug for future reference, and in some instances this should be carried on his person.

In severe forms of serum sickness type reactions, this regimen is suggested. (1) Discontinue the drug. (2) Diphenhydramine, 25 to 50 mg. every 4 to 6 hours (primarily for relief of itching and sedation). (3) Local lotions and soothing baths if the skin is involved. Zinc-shake lotion or calamine lotion applied with a brush, colloid baths (Aveeno), or wet dressings will help relieve pruritus and other skin complaints. (4) If needed, corticosteroids such as prednisone, 5 to 10 mg. every 4 to 6 hours, with gradual withdrawal over a 10-day period when the patient improves. With benzathine penicillin, longer therapy may be needed. (5) Occasionally, with rapidly developing wheals and pruritus, epinephrine, 0.2 to 0.4 ml. of 1:1000 solution subcutaneously, may be helpful until other measures can be applied. (6) Joint symptoms can frequently be helped with salicylates.

In the past, penicillinase, an enzyme of bacterial origin, has been used in treatment of penicillin reactions of the delayed type. This enzyme converts penicillins into penicilloic acid and was used with the hope that it would rapidly remove the supposed offending hapten. This might be

the case with free unbound penicillin,[23] but the repository salts of penicillin such as benzathine penicillin and procaine penicillin would not be affected, since they are not in solution and as a result are unavailable for hydrolysis. It is also of interest to note that penicilloic acid has been considered to be one of the precursors of the minor determinants.[14] There is no evidence that penicillinase has any effect on any of the major or minor determinants discussed previously. If these substances are the antigenic material that is responsible for penicillin allergy, penicillinase would not be expected to have any effect on the clinical course of a patient with the disorder.

Penicillinase would not seem to be useful against the sodium and potassium salts of penicillin since these soluble salts are cleared very rapidly by the kidney under normal circumstances. The cephalosporins and the penicillinase-resistant penicillins such as methicillin are not hydrolyzed by the penicillinase in current use, and the use of the enzyme in situations involving these substances would be illogical.

Penicillinase, because of its protein nature has caused severe allergic reactions.[23, 48]

Anaphylactoid Reactions. Therapy of this complication must be planned in advance. Every physician who administers drugs must be prepared for and have the materials needed for the treatment of this dreaded complication available at all times. In a hospital situation this is usually the case, but may not be so in home or office visits. The following procedure should be followed as soon as an indication of a reaction appears. (1) Place the patient in a supine position. (2) Give epinephrine, 0.2 to 0.4 ml. of 1:1000 solution, intramuscularly. If pulse or blood pressure is unobtainable, use 0.1 to 0.2 ml. intravenously. Repeat in 5 to 10 minutes or as needed. (3) Oxygen by mask. (4) Intravenous administration of 5 per cent dextrose in water. (5) Diphenhydramine, 50 mg. intravenously. (6) Hydrocortisone succinate, 200 to 300 mg. (7) For wheezing give aminophyllin, 500 mg. in 250 ml. of 5 per cent dextrose in water over 1 to 2 hours. (8) If blood pressure remains abnormally low, large amounts of fluids may be needed, with monitoring of the central venous pressure. Vasopressors such as norepinephrine or metaraminol tartrate are sometimes indicated.

ANTIBIOTIC THERAPY IN THE PATIENT WITH A PREVIOUS HISTORY OF AN ALLERGIC REACTION

In most situations in which a history of allergy to an antibiotic is elicited, the large number of available antibiotics will allow the physician to prescribe an alternate drug with little, if any, penalty to the patient. In a critical situation, the physician may elect to do one of three things. (1) Use an alternate antibacterial substance and accept the therapeutic penalty. (2) Use the antibiotic of choice and treat any allergic manifestations. (3) Desensitize the patient to the antibiotic.

1. *Alternate drugs.* In the main, the problem of suitable alternate drugs becomes critical with the penicillins. In the case of gram-negative

infections there are usually several drugs that are available, such as ampicillin, cephalothin, kanamycin, chloramphenicol, and the tetracyclines. With penicillin, the problem can be more serious because of the clear advantages of this drug in certain life-threatening infections. Most authors writing about therapy recognize this prevalent problem and give advice on alternate compounds (see Table 1).

Two less common infections deserve special mention. Subacute bacterial endocarditis in a penicillin-sensitive patient is a difficult clinical problem. At present, it would appear that the best alternate drug for staphylococcal endocarditis and endocarditis due to Streptococcus viridans is vancomycin.[21, 63] Vancomycin is bactericidal and this makes it appear to be superior to the bacteriostatic drugs such as erythromycin in this situation. Vancomycin is also a relatively toxic drug and must be used with discretion.

Another rare infection that may be encountered in a penicillin-sensitive person is tetanus. While clinical data are lacking, there is experimental evidence that 2 gm. of tetracycline per day may give comparable results to penicillin in therapy or as prophylaxis.[27]

2. *Desensitization.* Desensitization to penicillin and other antibacterial substances has been used successfully on occasion.[63] It consists of gradually neutralizing available antibodies by the administration of small, nonreacting doses, with a gradual increase in the amount

Table 1. *Alternate Antibacterial Therapy in Penicillin-Sensitive Patients*

DISEASE	ALTERNATE ANTIBIOTIC	DOSE
Pneumococcal pneumonia	Tetracycline	500 mg. every 6 hours[a, b]
	or	
	Erythromycin	500 mg. every 6 hours
β-Hemolytic streptococcal infection		
Therapeutic	Erythromycin	250 mg. 4 times a day
Prophylactic	Sulfadiazine	0.5 gm. 4 times a day[c]
Meningococcal meningitis	Tetracycline +	500 mg. every 6 hours[a, b]
	Sulfadiazine	1 gm. every 4 hours[d]
Staphylococcal infection	Lincomycin	500 mg. 3 times a day
	Vancomycin[f]	1 gm. every 12 hours[e]
Gonococcal urethritis	Erythromycin	500 mg. 4 times a day for 5 days
Syphilis	Erythromycin	500 mg. 3 times a day for 10 to 12 days

[a]Oral dose.
[b]Intramuscular dose is 250 mg. every 12 hours (less with impaired renal function).
[c]After another drug to eliminate the bacteria.
[d]After an initial loading dose of 4 gm.
[e]With normal renal function.
[f]In serious infections only.

given until full doses are reached. This procedure is obviously risky and difficult to evaluate, especially with penicillin. As noted earlier, the incidence of penicillin reaction in known penicillin-sensitive patients is only about 29 per cent, so that any small series of reported successful desensitizations must be viewed with this in mind.

3. *Steroids and Antihistamines.* Raper and Kemp[41] have reviewed the use of steroids in penicillin-sensitive patients with endocarditis and concluded that it was justified in some instances. It was clear from their case material that it was possible to suppress emergent symptoms of penicillin hypersensitivity, and in the cases reported by them, no immediate reactions occurred. The dose of steroids (as prednisone) was usually begun at 60 mg. per day and reduced gradually to a level sufficient to keep the patient free of allergic symptoms.

In general, the use of steroids simultaneously with antibiotics in infectious disorders is risky because of the masking of symptoms and signs by the steroids. It is clear that only in the most serious illness, and where an alternate antibiotic is not feasible, should it be considered. In the authors' opinion, the use of steroids and desensitization should be used only on the rarest of occasions, and only if an alternate drug is not feasible or has been used and failed.

ACKNOWLEDGMENTS

The authors wish to thank Miss Judy Harris and Kenneth Olschansky for their generous help with the manuscript and A. J. Kasik for his help with the references.

REFERENCES

1. Austen, K. F.: Systemic anaphylaxis in man. J.A.M.A., *192*:108, 1965.
2. Baer, R. I., and Harber, L. C.: Photosensitivity induced by drugs. J.A.M.A., *192*:989, 1965.
3. Batchelor, F. R., Dewdney, J. M., and Gazzord, D.: Penicillin allergy: The formation of the penicilloyl determinant. Nature, *206*:362, 1965.
4. Batchelor, F. R., Dewdney, J. M., Feinberg, J. G., and Weston, R. D.: A penicilloylated protein impurity as a source of allergy to benzylpenicillin and 6-aminopenicillanic acid. Lancet, *1*:1175, 1967.
5. Berté, S. J., DiMase, J. D., and Christianson, C. S.: Isoniazid, para-aminosalicylic acid, and streptomycin intolerance in 1,744 patients. Amer. Rev. Resp. Dis., *90*:598, 1964.
6. Brayton, R. G., MacLean, R. A., and Louria, D. B.: Cephaloridine: Clinical trial in gram-positive and gram-negative infections. Antimicrobial Agents and Chemotherapy, 1966, pp. 96-100.
7. Brown, W. J., Simpson, W. G., and Price, E. V.: Reevaluation of reactors to penicillin in V.D. clinic patients. Public Health Reports, *76*:189, 1961.
8. Carney, R. G.: Topical use of antibiotics. J.A.M.A., *186*:124, 1963.
9. Catterall, R. D.: Collagen disease and the chronic biological false positive phenomenon. Quart, J. Med., *30*:41, 1965.
10. Cieco, M. H.: Cross-allergenicity of the penicillins and cephalosporins. Arch. Intern. Med., *119*:141, 1967.
11. Cluff, L. E., and Johnson, J. E.: Drug fever. Progress in Allergy, *8*:144, 1964.
12. Cozart, A. G.: Letter to editor. J.A.M.A., *192*:227, 1965.
13. Derbes, V. J.: The fixed eruption. J.A.M.A., *190*:765, 1964.
14. DeWeck, A. L., and Schneider, C. H.: Penicillin antigens. *In* Rose, B., ed.: Allergology. Amsterdam, Excepta Medica, 1967, pp. 193-202.
15. DeWeck, A. L., Schneider, C. H., and Gutersohn, J.: The role of penicilloylated protein, impurities, polymers and dimers in penicillin allergy. Int. Arch. Allergy, *33*:535, 1968.
16. Editorial: Allergy to penicillin. J.A.M.A., *160*:778, 1956.

17. Editorial: Alternatives to penicillins, erythromycin group. Brit. Med. J. *1*:1213, 1963.
18. Editorial: Cross-allergenicity of penicillin and cephalosporin. J.A.M.A., *199*:495, 1967.
19. Eisen, H.: Hypersensitivity to simple chemicals. *In* Lawrence, H. S., ed.: Cellular and Humoral Aspects of the Hypersensitivity States. New York, Hoeber Medical Division, Harper & Row, 1959, p. 89.
20. Garrod, L. P.: Sensitivity reactions to antibiotics. *In* Rosenheim, M. L., and Moulton, R., eds.: Sensitivity Reations to Drugs. Oxford, Blackwell Scientific Publications, 1968, pp. 109-115.
21. Geraci, J. E., Nichols, D. R., and Wellman, W. E.: Vancomycin in serious staphylococcal infection. Arch. Int. Med., *109*:507, 1962.
22. Gilbert, F. I.: Cholestatic hepatitis caused by esters of erythromycin and oleandomycin. J.A.M.A., *182*:1048, 1962.
23. Greaves, A. B.: Further observations on penicillin. J.A.M.A., *176*:951, 1961.
24. Griffith, R. S., and Block, H. R.: Cephalothin, a new antibiotic: Preliminary clinical and laboratory studies. J.A.M.A., *189*:823, 1964.
25. Johnson, D. F., and Hall, W. H.: Allergic hepatitis caused by propionyl erythromycin ester of lauryl sulfate. New Eng. J. Med., *265*:1200, 1961.
26. Kershbaum, B. A., and Beerman, H.: Photosensitivity due to drugs. J. Med. Sci., *248*:445, 1964.
27. Kiser, J. S., de Demello, G. C., Reichard, D. H., and Williams, J. H.: Chemotherapy of experimental clostridial infections. J. Infect. Dis., *90*:76, 1952.
28. Kirby, W. M. M.: Clinical status of cephalothin. *In* Antimicrobial Agents and Chemotherapy, 1964, pp. 274-279.
29. Kirshbaum, B. A., and Beerman, H.: Photosensitivity due to drugs. J. Med. Sc., *248*:445, 1964.
30. Knudson, E. T., Robinson, O. P. W., Croydon, E. A. P., and Tees, E. C.: Cutaneous sensitivity to purified benzylpenicillin. Lancet, *1*:1184, 1967.
31. Lee, S. L., Revero, I., and Siegel, M.: Activation of systemic lupus erythematosus by drugs. Arch. Int. Med., *117*:620, 1966.
32. Levine, B. B.: Immunochemical mechanism of drug allergy. Annual Rev. Med., *17*:23, 1966.
33. Lockhart, W. E.: Letter to the editor. J.A.M.A., *192*:227, 1965.
34. Meyler, L.: Side Effects of Drugs. New York, Excerpta Medica Foundation, 1966, pp. 284-341.
35. Minkin, W., and Lynch, P. J.: Incidence of immediate systemic penicillin reactions. Military Med., *133*:557, 1968.
36. Nesmith, L. W.: Hemolytic anemia due to penicillin. Report of a case in which antipenicillin antibodies crossreacted with cephalothin sodium. J.A.M.A., *203*:27, 1968.
37. Perlman, F.: Letter to the editor. J.A.M.A., *192*:64, 1965.
38. Pirila, V., and Pirila, L.: Sensitization to the neomycin group of antibiotics. Acta Derm.-Venereol., *46*:489, 1966.
39. Pirila, V., Hirvonen, M. L., and Rouhunkaski, S.: The pattern of cross-sensitivity to neomycin. Dermatologica, *136*:321, 1968.
40. Raffel, S.: Immunity. 2nd ed. New York, Appleton-Century-Crofts, 1961, p. 361.
41. Raper, A. J., and Kemp, V. E.: Use of steroids in penicillin-sensitive patients with bacterial endocarditis. New Eng. J. Med., *273*:297, 1965.
42. Rees, R. B.: Cutaneous reactions to antibiotics. J.A.M.A., *189*:685, 1964.
43. Reidenberg, M. M.: Registry of adverse drug reactions. J.A.M.A., *203*:31, 1968.
44. Rosenthal, A.: Follow-up study of fatal penicillin reactions. Special report. J.A.M.A., *167*:1118, 1958.
45. Simon, H. J., and Rantz, L. A.: Untoward reactions to antibiotics. Annual Rev. Med., *12*:111, 1961.
46. Smith, J. W., Johnson, J. E., and Cluff, L. E.: Studies on the epidemiology of adverse drug reactions. II. An evaluation of penicillin allergy. New Eng. J. Med., *274*:998, 1966.
47. Smith, J. W., Seidl, L. G., and Cluff, L. E.: Studies on the epidemiology of adverse drug reactions. V. Clinical factors influencing susceptibility. Ann. Intern. Med., *65*:629, 1966.
48. Smith, T. J.: Antibiotics and induced disease. *In* Moser, R. H., ed.: Diseases of Medical Progress. 2nd ed. Springfield, Illinois, Charles C Thomas, 1964, pp. 3-60.
49. Smith, V. M.: Fatal reactions to penicillin. New Eng. J. Med., *257*:447, 1957.
50. Steinman, H., and Brandriss, M.: Current developments in penicillin therapy and hypersensitivity. Med. Ann. D.C., *31*:689, 1962.
51. Stewart, G. T.: Allergic residues in penicillin. Lancet, *1*:1177, 1967.
52. Stewart, G. T.: Macromolecular residues contributing to allergenicity of penicillins and cephalosporins. *In* Antimicrobial Agents and Chemotherapy, 1967, pg. 543.
53. Stewart, G. T.: Allergic factors in penicillins and cephalosporins. Amer. Heart J., *75*:429, 1968.
54. Symmers, W. St.C.: The occurrence of the so-called collagen diseases and of other diseases systemically affecting the connective tissues as a manifestation of sensitivity to

drugs. *In* Rosenheim, M. L., and Moulton, R., eds.: Sensitivity Reactions to Drugs. Oxford, Blackwell Scientific Publications, 1958, pp. 209-222.
55. Thoburn, R., Johnson, J. E., and Cluff, L. E.: Studies on the epidemiology of adverse drug reactions. IV. The relationship of cephalothin and penicillin allergy. J.A.M.A., *198*:345, 1966.
56. Tozer, F. L., and Kasik, J. E.: The medical-legal aspects of adverse drug reactions. J. Clin. Pharm., *8*:637, 1967.
57. Triacetyloleandomycin. Med. Letter on Drugs and Therapeutics, *4*:6, 1962.
58. VanArsdel, P. P.: Allergic reactions to penicillin. J.A.M.A., *191*:238, 1965.
59. VanArsdel, P. P., and Gilliland, B. C.: Anemia secondary to penicillin treatment. J. Lab. Clin. Med., *65*:277, 1965.
60. Welch, H. C. N., Weinstein, H. I., and Boechman, B. B.: Severe reactions to antibiotics. Antibiotic Med. and Clinical Therapy, *4*:800, 1957.
61. Weston, R. D.: Penicilloylated protein contaminating 6-APA and benzylpenicillin. *In* Antimicrobial Agents and Chemotherapy, 1967, pp. 553-559.
62. White, J. M.: Penicillin induced hemolytic anemia. Brit. Med. J., *3*:26, 1968.
63. Winterbauer, R. H., Ronald, A. R., Belcher, D. W., and Turck, M.: Antimicrobial therapy in patients sensitive to penicillin. J. Chron. Dis., *20*:407, 1967.

950 East 59th Street
Chicago, Illinois 60637

Kidney Transplantation: Patient Selection and Management

ADRIAN I. KATZ, M.D.*

In less than a decade, the prognosis for the patient with chronic uremia has improved considerably as a result of the development of chronic hemodialysis and kidney transplantation as practical therapeutic modalities. Whereas only 10 years ago there was little one could do for the patient with far advanced renal failure besides palliation and symptomatic treatment of the various manifestations of his disease, today, maintenance dialysis and kidney grafting offer not only the means of prolonging life, but, to an increasing number of individuals, also the possibility of rehabilitation as productive members of society. Unfortunately, the demand for these facilities so outstrips their availability that only a fraction of the patients who in theory could benefit from them, actually do. In this country, fewer than 1000 out of the 8000 patients who qualify for it each year have access to a dialysis-transplantation program. While the limited availability of dialysis is due chiefly to lack of equipment and trained personnel, that of transplantation is caused mainly by the currently insufficient supply of donor kidneys.

Kidney transplantation encompasses a host of ethical, moral, and legal problems in addition to those of organ procurement and the medical problems involved in grafting of organs between genetically dissimilar individuals. Despite these difficulties, the last decade has witnessed a steady improvement in the outcome of kidney transplantation (Fig. 1), the result of the combined efforts of the surgeons, internists, urologists, and immunologists involved in this endeavor of major biological as well as practical significance. This presentation will review briefly the present status of kidney transplantation and the practices currently in use for selection and management of the transplant recipient.

DIALYSIS OR TRANSPLANTATION?

In most centers, chronic dialysis and kidney transplantation are done today in conjunction with each other. Obviously, no transplantation

*Assistant Professor of Medicine, University of Chicago

Supported in part by U.S. Public Health Service General Research Support Grant FR-5367 and the Chicago Heart Association.

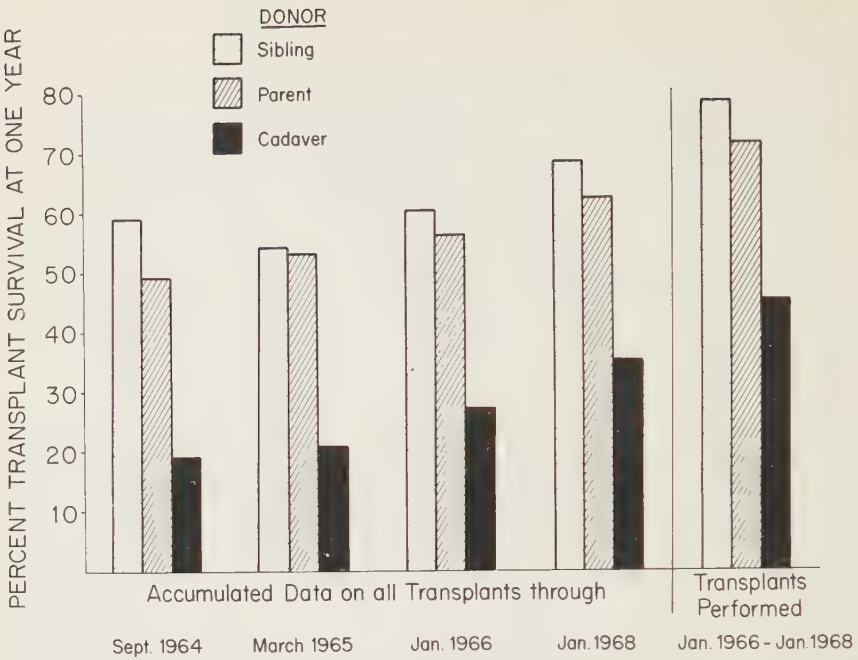


Figure 1. Worldwide survival statistics of patients with kidneys obtained from siblings, parents, and cadaver donors, summarized from the Kidney Transplant Registry report.[36] Note steady improvement in survival of grafts from all donor categories.

program can operate without the support of dialysis, which is required both to prepare the patient for surgery and to support him until the graft is fully functional, or if it is rejected, until another kidney is found. Although one can maintain the terminal uremic patient alive on chronic dialysis for periods of up to several years, and average length of survival appears to be greater on dialysis than after transplantation,[20] more and more centers use the dialysis program in support of transplantation rather than as an end in itself, for two main reasons. First, without an open end to remove some of the patients, any dialysis program, regardless of its size, will soon become filled and any further intake of patients will be impossible. Secondly, and more important, the shift in emphasis from strictly chronic dialysis to dialysis-transplantation programs was probably due to the recognition that transplantation, if successful, has more to offer the chronic uremic patient than dialysis as such. The increasing number of institutions at which transplantation can be performed competently and the continuously improving survival statistics of transplanted patients have greatly enhanced this trend.

Survival data of 302 patients maintained on chronic dialysis recently compiled from 14 centers[22] show 87 per cent survival at 1 year and 77.3 per cent at 2 years, with a steady but moderate decline thereafter. The projected 7 years' survival was 57.8 per cent. It is evident that many patients whose life expectancy would have been otherwise measured in

months, can be kept alive today by chronic dialysis for long periods of time. Although it can prolong life, there are, however, serious limitations to what maintenance dialysis can accomplish. While it partially substitutes for the loss of the excretory capacity of the kidney, dialysis does not replace its nonexcretory functions, such as those concerned with the homeostasis of calcium and phosphorus, the formation of erythropoietin, or the regulation of the arterial blood pressure. Consequently, many of the complications of uremia are not improved by dialysis, and some may be made worse by it. To name but a few, these are anemia, hypertension, peripheral neuropathy, renal osteodystrophy, impaired potency and fertility, and the inability of uremic children to grow and mature sexually.[17, 23, 31, 38]

Unless dialysis is performed overnight in his home, the chronic dialysis patient who has to be connected to the artificial kidney for many hours two or more times a week has both his freedom and the possibility of regular employment seriously limited. In addition, his salt and water intake are restricted in order to prevent excessive accumulation between dialyses. Successful kidney transplantation, on the other hand, may return the patient to near normal health and permit regular activity and employment. It should be emphasized, however, that by a successful transplant is meant not only a well functioning graft, but also that the immunosuppressive therapy required for its survival does not affect the recipient's well being and normal activities. Despite the remarkable progress achieved in kidney transplantation in the past few years, the number of patients who have achieved this is not satisfactory yet.

There is perhaps no need, however, to compare dialysis and transplantation, since they should be used not as alternative, but rather as complementary therapeutic approaches. Indeed, it appears that the interests of the severely uremic patient are best served when maintenance dialysis and transplantation are both available and are undertaken together as part of his management.

SELECTION OF CANDIDATES FOR TRANSPLANTATION

The chronic uremic patient should be managed conservatively as long as he remains comfortable and active. Every attempt should be made to treat superimposed reversible conditions such as dehydration, salt depletion, urinary tract obstruction, and infection. Since it is not always fully appreciated that the ability of the chronically diseased kidney to conserve sodium and water is markedly impaired, unnecessarily drastic salt restriction, particularly in conjunction with diuretics, or a more protracted episode of vomiting or diarrhea without adequate replacement, may result in significant salt and water deficits, which will further compromise renal perfusion. A useful guide to the patient's daily salt requirement is provided by the amount excreted in the urine over 24 hours, or, when this determination is not available, by the fluctuations in body weight. Infection, which is often found on morphological examination to be superimposed on the underlying lesion, may also place an additional burden on the diseased kidney.

The correction of any of the conditions mentioned above, although not affecting the underlying disease, may result in sufficient improvement to allow months, and occasionally several years, of relative well being and temporarily delay the need for dialysis. The use of a diet low in protein but including all the essential amino acids[7] is likewise beneficial in some chronic uremic patients, particularly if there is some residual renal function left (creatinine clearance above 3 ml. per min.). On the other hand, it cannot be emphasized strongly enough that undue procrastination in initiating dialysis when it becomes necessary may permit the development of complications which will jeopardize both the subsequent management on dialysis and the chances of successful transplantation. Prominent among these are peripheral neuropathy, pericarditis, malnutrition, serious electrolyte disorders, and renal osteodystrophy.

Some of these complications, in particular renal osteodystrophy and peripheral neuropathy of the motor variety, if allowed to progress, are unlikely to be corrected by dialysis and sometimes not even by a well functioning kidney graft.[11, 21] Since considerable delay is often caused by logistic considerations, such as availability of a place on the dialysis program and the time required for complete medical evaluation of the patient and his potential donors, as well as for obtaining financial support for the extremely costly dialysis and transplantation, preliminary evaluation of the candidate and initial enquiries about potential donors are never too early. Approaching the patient and the family earlier in the course of the disease will also allow them to cope better with the considerable personal, emotional, and financial adjustments that the situation will demand of them.

The criteria for selection of patients as transplant candidates vary from center to center, and change frequently in parallel to changes in the results of transplantation and the availability of organs. Since most patients have to undergo maintenance dialysis for varying periods prior to transplantation, individuals considered for transplantation should be suitable for chronic dialysis as well. The medical criteria for acceptance on any of the two programs are generally similar. However, patients who will have to be maintained on dialysis for longer periods (particularly those who lack a living donor and will have to wait for a cadaver kidney) should be mature, emotionally stable, and reasonably intelligent individuals in order to insure their full cooperation during the long and stressful wait for a graft.

There are several basic medical criteria that are required by most programs for acceptance of an individual as a candidate for transplantation. First, he should have exhausted the benefits of the conservative measures of treatment available to maintain a useful existence. Second is the consideration of age. There is no lower age limit: kidneys have been transplanted successfully in children, even as young as 3 years of age.[40] The question which most often arises in relation to age is the upper limit above which transplantation is not considered justified. Although there are no rigid rules, at present most centers will usually not accept candidates above 50 years of age.[11] There is, however, no compelling reason to

deny a patient who would otherwise be a suitable candidate the chance of receiving a kidney graft only because he is past an arbitrary chronologic age limit. One should insure, though, that such a patient is free of overt cerebrovascular and particularly coronary heart disease and that his general health (exclusive of his kidney failure) is adequate for at least 5 years of life expectancy should the transplantation succeed.[26] Transplantation is contraindicated at any age if the recipient has an irreversible disease associated with a fatal outcome, such as severe brain or liver damage, or widespread cancer.

An important consideration in the choice of candidates for transplantation is whether one deals with a primary renal condition or with involvement of the kidney in the course of a systemic disease, such as disseminated lupus erythematosus, periarteritis nodosa, or diabetes mellitus. At present, in most centers, these patients with secondary renal involvement are not accepted as candidates, probably not so much because of the chance of recurrence of the original disease in the transplanted kidney—which in the case of diabetes may take longer than the life expectancy of the graft—but because of the effects of the primary disease on organs other than the kidney. Although grafting a kidney from a living donor in a patient with a serious systemic disease is not justified, such an individual might nevertheless also enjoy the additional life expectancy conferred by a functioning kidney graft should an adequate supply of cadaver kidneys become available in the future. As pointed out by Merrill,[26] the immunosuppressive therapy employed to protect the graft from rejection, may also protect it from the effects of the systemic disease.

A normal lower urinary tract into which to implant the new kidney is required by most transplantation surgeons. As is true of several other criteria, this requirement is also not absolute, as transplantation has been successfully performed with the ureter implanted into an ileal-loop conduit serving as an artificial bladder. Finally, transplantation should be postponed until such conditions as active glomerulonephritis, infection, leukopenia, active peptic ulcer, or severe malnutrition subside or can be corrected.

In summary, it appears that the only absolute requirements which a patient with renal failure has to fulfill in order to become a candidate for transplantation are those of having exhausted all other therapeutic modalities available and of being free of an irreversible and potentially fatal nonrenal disease. All the other criteria currently in use vary from one center to another and are determined by the size of the program, the proficiency of its surgeons and urologists, and in particular by the availability of donor kidneys. It is likely that if cadaver organ procurement and the efficiency of matching the recipient with the best available donor will improve significantly, many patients who are denied kidney transplantation today, will be able to avail themselves of this opportunity in the future.

SELECTION OF KIDNEY DONORS

The three main sources of kidneys are from parents, siblings, and cadaver donors, if one excludes the unrelated living donors, a category which is probably no longer justified today. Figure 2 shows the most

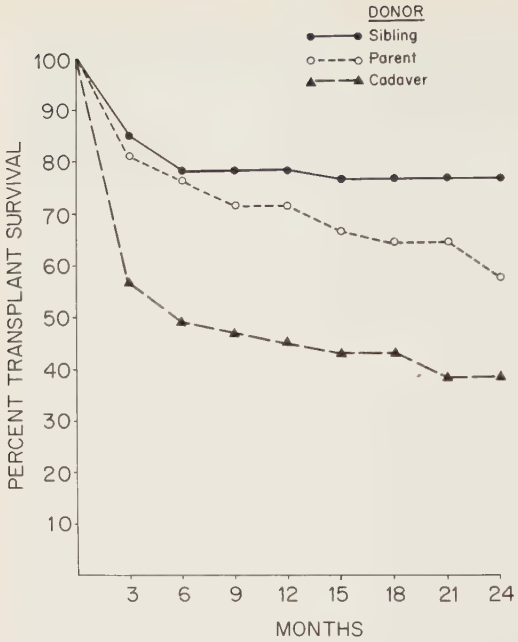


Figure 2. Transplant survival curves during the first 2 years after transplantation, modified from the Kidney Transplant Registry Report.[36] The steep initial decline subsequently levels off, a feature particularly evident in recipients of kidneys from siblings.

recent transplant survival data* obtained from 736 transplants performed since January 1966.[36] It is evident that survival of grafts from siblings is longer than of those from parents, and that both last longer than a kidney obtained from a cadaver. The survival curves characteristically show a rather steep initial decline which slows down thereafter. This feature can be better understood if one appreciates that the central problems encountered in transplantation, with the rare exceptions of those between identical twins, are due either to failure of the graft or to complications of the therapy employed to prevent or overcome its rejection. Since these usually occur relatively early after transplantation, the risk of graft deterioration decreases with time.[14] The emerging experience indicates that the "cutoff point" beyond which rejection becomes less frequent is 6 to 12 months for recipients of a kidney from a sibling, and 1 to 2 years for those receiving the kidney of a parent or a cadaver. Although later deterioration of the allograft sometimes does occur, it is the sharp decrease in the frequency of rejection beyond this "cutoff point" which accounts for the shape of the survival curves.

Provided there is a choice of donors, which is rarely the case, the selection should be based primarily on the expected chance of the transplanted kidney to survive.† Although a rough estimate of what category of donor kidney among those available for a particular recipient will fare best can be derived from the figures accumulated to date from

*Patient survival, not reviewed here, is much better, since after failure of the graft and its removal, the patient is returned to maintenance dialysis and subsequently may receive a second or even a third or fourth transplant.

†Very often, however, other considerations have to be applied. For example, although siblings are better donors than their parents, their motivation to donate a kidney is nearly always less sound, or they are too young to be accepted as donors.

a large number of transplantations, it should be appreciated that this information is meaningful only on a statistical basis, and can be useless for any individual case. A more accurate, although by no means precise, estimate of the future transplant's fate is provided by the various techniques which test the histocompatibility between the recipient and his prospective donors.[1, 34] These are discussed in detail elsewhere in this issue. In general, if the donor and recipient match very well, the chances of graft survival are good, and if there are several major incompatibilities, they are poor.[29, 34, 43] Nevertheless, there are sufficiently frequent exceptions to this statement, particularly in poorly matched recipients who occasionally do surprisingly well, for some investigators to feel that poor histocompatibility should not exclude a potential donor.[26] There is general agreement, however, that if circulating cytotoxic antibodies against donor white blood cells are present in the recipient's serum the graft is likely to undergo massive, acute rejection.[19, 30, 45] Under these circumstances, transplantation is contraindicated. This test can be performed relatively quickly, so that even in the case of a cadaver donor, the potential recipients with circulating antibodies can be excluded.

Much of the controversy surrounding kidney transplantation, and in particular the moral and ethical points repeatedly raised,[6, 10] center on the propriety of removing an organ from a healthy individual and its implantation against considerable odds in a sick recipient. This consideration is underscored, in the view of some, by the feasibility of successful transplantation from cadavers. If it is possible, under proper circumstances, to obtain from a cadaver a functioning graft which has a reasonably good chance to survive, why should one subject a healthy donor to relatively major surgery and the unknown future risks which may accrue from removal of one kidney?

The answer to this question rests with the individual physician's experience and personal philosophy. Following are some of the author's reasons to believe that a living donor should be preferred whenever clinical or immunological evaluation by histocompatibility testing indicate that transplantation from that particular donor is likely to be successful:

First, the over-all experience is that survival of kidneys from related living donors is unquestionably better than for those obtained from cadavers, even though an occasional center achieves satisfactory results with the latter.[18] It is easy to understand why, with the relatively limited recipient pools that each center has at present, the likelihood of matching a given recipient with a randomly selected unrelated donor who happens to have died close to the transplantation center is quite small[42] and can rarely, if ever, be as good as the best available match from the recipient's own family. Moreover, there are significant differences in the postoperative course of recipients of a kidney from a living relative as compared to a cadaver donor, which are not always reflected in the survival statistics. Because of the longer period of ischemia, cadaver kidneys, even if not rejected, frequently undergo a protracted period of acute tubular necrosis in the recipient, an event which occurs only rarely in recipients of a kidney from a well matched living donor. The period of anuria is usually longer in these cases than in the acute tubular necrosis

which occurs in a previously healthy individual in his own kidneys, and the chances of intercurrent infection are greatly enhanced by the massive immunosuppressive therapy required at this stage. The difficulty of detecting rejection and of regulating the dose of immunosuppressive drugs in the absence of adequate excretory function often results in bone marrow suppression or severe liver damage during this phase.

Secondly, since prospective matching with potential living donors can be done at leisure, the best possible donor in the light of the information so obtained can be selected. In addition, transplantation can be scheduled as early as necessary, thus eliminating the long wait for an adequate cadaver kidney, during which many patients deteriorate and some may die even though treated by dialysis.

A final argument in favor of living donor transplantation is provided by evaluation of the immediate as well as the long-term risks to the donor involved. Although these dangers, and in particular the unknown long-term ones, should not be minimized, the experience accumulated so far with approximately 1000 living donor transplants, is encouraging enough to justify their continuation.

The immediate operative mortality risk of unilateral nephrectomy in a healthy individual is 0.05 per cent;[10] the long-term risk, more difficult to assess, can be estimated by actuarial methods based on the long experience available with unilateral nephrectomy. Evidently, diffuse parenchymatous diseases have no more tendency to affect one kidney than when both are present, and when the involvement is diffuse it matters little whether there are one or two kidneys available. The risk question does become important in those conditions in which one kidney may be involved by stones, trauma, or neoplasms, which may require its removal. The calculated long-term risks from these conditions is 0.07 per cent, and the total risk incurred to the donor is therefore 0.12 per cent.[10] Using a different terminology, the chance that a healthy 35 year old man will survive for the next 5 years drops from the normal of 99.3 to 99.1 per cent if he is a kidney donor, a risk equal to that of driving a car 8 miles to work and back every day.[25] Arguments for or against living donor transplantation should be evaluated in light of this arithmetic.

Although no fatality has been reported so far in a kidney donor as a result of his gift, it is conceivable that this will happen in the future. Since there is a calculated risk, albeit small, which the donor has to take, he should be fully informed both of the dangers involved and the uncertainty of success of the transplant. In addition he should be given all the time necessary to reflect and should not only be clearly informed that he is free to withdraw from consideration, but also that the physician will supply him with a medical explanation for such withdrawal if he so chooses. Because of the tremendous pressures and agonizing decisions that often have to be faced, it is best to leave to the family the task of producing the volunteers among whom the donor will be selected. The physicians in charge of the recipient should refrain from any persuasion or solicitation, and limit themselves to discussing the statistical chances of success of various donors, including cadavers, in view of the current experience.

What is required of a kidney donor? He should be a well motivated and emotionally balanced individual in good general health, who, besides having ABO (but not Rh) compatibility with the recipient, is a good enough match, as assessed by tissue typing, to justify the use of one of his kidneys. Although there are several donors in their late 60's and early 70's on record, younger donors should be preferred since they are less likely to have significant atherosclerosis of the larger vessels or nephrosclerosis. They should be, however, past 21, the age of legal consent. Preliminary screening of prospective donors requires only a blood sample on which histocompatibility testing and red cell cross matching are done. The remainder of the examination for the compatible donor, which is done in the hospital, includes, in addition to a complete history, physical and laboratory examination, and a thorough psychiatric evaluation, tests specifically directed to evaluate renal function and anatomy (creatinine clearance, urinalysis, and urine cultures, intravenous pyelogram and bilateral selective renal arteriograms). In order to be accepted as a donor, the candidate should be free of high blood pressure, kidney disease, or any systemic illness that may contraindicate anesthesia or surgery. In addition, since multiple renal arteries bilaterally increase the technical difficulties considerably, most transplantation surgeons prefer to deal at least on one side with a single renal artery.

If a living related donor is not available, a situation more commonly encountered than not, the candidate has to wait for a suitable cadaver donor. Cadaver donors are selected from patients dying of head injuries or other irreversible brain lesions, and, less commonly, after heart surgery. Patients with cancer (other than intracerebral), septicemia, renal disease, and prolonged shock or oliguria preceding death are excluded. A urinalysis and serum creatinine test may be helpful to screen out inadequate donors, and should be obtained if time permits. As in the case of the living donor, blood should be obtained for cross match and histocompatibility testing with the potential recipient.

In addition to the psychological tensions generated by the uncertainty as to the date of transplantation and the possible deterioration during the waiting period, the results of cadaver kidney transplantation are inferior to those obtained from living donors (Figs. 1 and 2). Lately, however, the exclusion of donors in whom the circulation was compromised before death, and reduction of the ischemic time between death and implantation, have resulted in significantly improved results. If preservation of the kidney for periods sufficiently long to allow prospective tissue typing and selection of the best possible recipient should become available,[2] the results of cadaver kidney transplantation may improve further and approach those currently obtained with living donors.

MANAGEMENT OF THE TRANSPLANT RECIPIENT

Preparation for Transplantation

If there is a living donor, transplantation should be scheduled early, before the patient becomes malnourished or infected, or the irreversible complications of uremia have set in. Since the date of transplantation

for the cadaver kidney recipient cannot be predicted, he should be started on chronic dialysis as soon as it becomes apparent that the conservative management is not satisfactory; while on maintenance dialysis he becomes a candidate for transplantation should a well matched kidney become available.

It is desirable to initiate dialysis at least 4 to 6 weeks prior to the scheduled transplantation. During this period the patient becomes acquainted with the procedure, and existing fluid, electrolyte, or nutritional disorders can be improved or corrected and his own kidneys removed. In most centers the recipient's own kidneys are removed prior to transplantation, primarily because blood pressure on dialysis is better controlled if the patient is anephric. Nephrectomy may also remove a source of infection, or, in the case of polycystic kidney disease, pain and bleeding. Finally, the transplantation itself is better tolerated if it is not done together with bilateral nephrectomy at a time when the patient is receiving massive immunosuppressive therapy.

The presence of infection contraindicates transplantation, as it does any other major surgery. Leukopenia and active peptic ulcer are relative contraindications, since their presence will render the post-transplantation management with immunosuppressive drugs difficult or impossible. On several occasions patients with active or bleeding peptic ulcer underwent gastrectomy in preparation for transplantation and did well thereafter. As long as any of these conditions is present, transplantation should be postponed and the patient maintained on chronic dialysis. Likewise, because of the significant risk of subsequent development of glomerulonephritis in the graft, it is recommended that transplantation be withheld if active progressive glomerulonephritis is present.[26]

Following Transplantation

As pointed out above, the essential condition for successful organ transplantation between individuals who are not identical twins is the suppression of the immune response of the host against the allograft. Much of the progress achieved so far in transplantation has been the result of the steady improvement in the efficiency and safety of the methods of immunosuppression available. Nonetheless, it should be emphasized that, with the possible exception of antilymphocytic serum, the ideal immunosuppressive agent which will selectively inhibit the rejection process without affecting other systems is not yet available.

A large number of procedures and immunosuppressive agents, summarized in Table 1, have been tried so far with varying degrees of success. Some, like extracorporeal irradiation of the blood or thoracic duct drainage, are used irregularly in some centers as adjuncts to, or after the failure of, the more commonly used techniques of immune suppression. Others, of historical interest now, have been all but abandoned, sometimes even by their own proponents. A detailed discussion of their relative merits and drawbacks is beyond the scope of this presentation, and the interested reader is referred to several sources of further information.[8, 14, 26, 27] Following is a brief description of the methods currently in widespread use:

The two most commonly used immunosuppressive agents are azathioprine (Imuran), employed chiefly for primary immunosuppres-

Table 1. *Methods of Immunosuppression*

Immunosuppressive Agents

AGENT	PRESUMED MODE OF ACTION	AVERAGE DOSE	COMMON SIDE-EFFECTS*	REMARKS
Azathioprine (Imuran)	Inhibits proliferation of immunologically competent cells by interfering with DNA and RNA synthesis	4 mg./kg./day initially 2 to 3 mg./kg./day maintenance (oral)	Bone marrow depression Liver damage	Cornerstone of primary immunosuppressive therapy; universally used
Corticosteroids (Prednisone)	Anti-inflammatory Lymphocytolytic ?Block antigen-antibody interaction ?Inhibit antibody synthesis	2 to 3 mg./kg./day initially 10 to 40 mg./day maintenance† (oral)	Peptic ulcer Gastrointestinal bleeding Psychosis Osteoporosis Steroid diabetes Aseptic necrosis of bone	Treatment of choice for rejection In some centers, particularly if ALS is available, the corticosteroids are reserved only for treatment of rejection
Antilymphocytic serum (ALS) or purified antilymphocytic globulin (ALG)	Blocks delayed hypersensitivity and homograft immunity by a primary effect on the circulating lymphocyte; the basic mechanism of this effect is unknown, but does not appear to be dependent on lymphocyte depletion	Schedules and dosage not uniform, because of variable titers and potency of the serum; usually daily IM injections for 5 days prior to transplantation and for variable periods post transplantation	Intense pain and induration at injection site Low grade fever Hives Anaphylaxis (rarely)	Powerful immunosuppressive agent when used in conjunction with other measures; clinical experience still limited
Actinomycin C	Blocks the synthesis of messenger RNA	4 to 5 μg./kg./day, IV	Nausea Vomiting Bone marrow depression	Used in the past as adjunct in the treatment of rejection; largely discontinued today

(Table continues on following page.)

Table 1. *Methods of Immunosuppression* (Continued)

Immunosuppressive Agents

AGENT	PRESUMED MODE OF ACTION	AVERAGE DOSE	COMMON SIDE-EFFECTS	REMARKS
6-Mercaptopurine	Same as azathioprine	1.5 to 2 mg./kg./day	Same as azathioprine	Not in use at present since the development of its less toxic derivative, azathioprine
Azaserine	Inhibits nucleotide synthesis	8 to 10 mg./day, oral or IV	Jaundice Delirium Coma	Not in widespread use at present

Other Measures

	COMMENTS
Local graft irradiation	Used in cadaver kidney recipients. 150 R × 4 or 5, every other day, starting on the day of transplantation. Course may be later repeated as adjunct therapy for rejection.
Extracorporeal irradiation of blood	Used in some centers as adjunct therapy. Performed through the same indwelling cannula used for chronic hemodialysis. Blood is irradiated as it passes through a radioactive source before returning to the patient.
Thoracic duct lymph drainage	Used in some centers as adjunct therapy. Should be started 5 to 7 days prior to transplantation, therefore not applicable to cadaver kidney recipients. Effectiveness apparently correlated with number of lymphocytes removed. Cell-free lymph has to be reinfused IV to prevent fluid and protein depletion. Sometimes accompanied by local or systemic infection.
Total body irradiation	Largely discontinued since the development of immunosuppressive agents.
Thymectomy Splenectomy	Evidence that they improve graft survival in man is lacking. Both techniques abandoned today.

*All immunosuppressive agents predispose to infection.
†Dose varies according to the functional integrity of the graft.

sion, and the adrenocorticosteroids, usually prednisone, which are the drugs of choice for the treatment of rejection. Imuran is started several days (2 to 8) before the scheduled transplantation date in recipients of a living donor kidney and on the day of surgery in those of a cadaver kidney, usually in a daily dose of 4 mg. per kg., which is maintained for about a week following transplantation and then tapered to 2 to 3 mg. per kg. per day. This dose level must then be maintained for long periods of time, and in the light of the present knowledge, possibly indefinitely. Prednisone is usually started on the day of surgery (2 to 3 mg. per kg. per day) and maintained at this level also for about a week, following which it is gradually tapered according to the outcome of transplantation. If it has been successful, the dosage of prednisone can be reduced to 20 to 40 mg. per day within several weeks, and to even lower levels thereafter. Most groups prefer to keep all patients on a very small dose of prednisone (as low as 5 to 10 mg. per day) for long periods, although it is not entirely clear whether doses of this magnitude are effective. In some centers, cadaver kidney recipients also receive local irradiation of the graft every other day starting on the day of transplantation, for a total of four or five doses of 150 R each.

A very promising recent addition to the immunosuppressive agents available is the heterologous antilymphocytic serum (ALS), or antilymphocytic globulin (ALG), a purified fraction of the serum, consisting mostly of gamma G globulin (IgG). Introduced for human homotransplantation by Starzl and his colleagues,[41] it was found to be useful in prolonging homograft survival when given together with azathioprine and prednisone. Because it permitted lowering the dose of the standard immunosuppressive agents and therefore eliminated some of their side-effects, significantly improved patient survival was also obtained. In the first series of 20 homografts (the majority from living related donors) reported by this group,[39] there was only one accidental death during the follow-up of 13 months. Renal function, as estimated by creatinine clearance, was also better in this group than in comparable groups of patients from the same center who did not receive ALS. The slowly accumulating reports from other centers,[28, 44] although somewhat less enthusiastic than that of the Denver group, are encouraging.

ALS is not yet in widespread use, since at the time of this writing it is not commercially available and therefore can be used only in centers which have the facilities to produce it. Lymphocyte suspensions from human thymus, spleen, or lymph nodes, and particularly those obtained by draining the thoracic duct, are repeatedly injected in horses, which are later bled and the serum process to extract a more or less purified antilymphocytic gamma globulin, while undesirable antibodies like the red cell agglutinins are removed by absorption. Since the procedures for obtaining the ALS are not standardized, and the quality of the antibody produced may vary from horse to horse, or even in the same animal on different occasions, there is no standard schedule of administration either. The serum is given intramuscularly, in a volume of 2 to 10 ml. according to the antibody titer present. In recipients of living donor kidneys, it is usually given daily for about a week prior to trans-

plantation and for varying periods up to 4 months in decreasing doses thereafter. Cadaver kidney recipients receive the serum daily for several weeks, and with decreasing frequency for several more months after transplantation. The most common side-effects are pain and induration at the injection site and a low-grade fever. Anaphylactic reactions occur rarely and are generally mild and easily controlled.[4, 16]

During the first few days following transplantation the patient is maintained on reverse precautions, his urine output—which can be excessive at this stage if the graft functions well—carefully monitored and the fluid and electrolytes lost therein adequately replaced. If rejection or surgical complications do not occur, the patient is gradually allowed to move about the ward, the diet and fluid intake are liberalized, and he may be able to leave the hospital within 3 or 4 weeks from transplantation. It is desirable, however, that the patient should live within easy traveling distance to the center for at least 6 months, both for rapid access in case of emergency and for the frequent outpatient visits required during this period. He is first seen twice a week for about a month, then once a week for 2 months, and finally every 2 or 3 weeks for about 3 more months in the transplantation outpatient clinic, and is instructed to inform his physicians by phone of any rise in temperature, decrease in urine output or other symptoms which may signal rejection or infection.

At each visit a physical examination and urinalysis is done, and blood is drawn for BUN, serum creatinine, and a complete blood count, and occasionally for serum electrolytes, LDH and GOT, or special tests. A 24 hour creatinine clearance is performed on the urine collected by the patient the previous day. The dosage of both azathioprine and prednisone is adjusted according to the results of kidney function tests and the presence of significant side-effects. Patients who do exceptionally well (normal renal function, normal urine sediment) may be maintained on only 1.5 mg. per kg. per day of azathioprine and a minimal dose of prednisone.

Management of Complications

Graft rejection, which may occur at any time after transplantation, is manifested by a sharp decline in urine output, rising BUN and serum creatinine levels, rising blood pressure, fever, and local tenderness at the operation site. These indices of rejection may appear in various combinations, or in association with more subtle laboratory changes,[3] but only rarely will all appear together. Careful observation of the recipient and frequent laboratory determinations should enable the experienced clinician to diagnose rejection accurately in the majority of cases, except possibly when it occurs during the acute tubular necrosis which frequently follows cadaver kidney transplantation. If in these circumstances rejection is suspected, its diagnosis may not be possible without a kidney biopsy. If the index of suspicion is high, it is probably wise, however, to treat the patient promptly, since by the time the results of the biopsy become available, the rejection process may no longer be reversible.

Acute tubular necrosis is managed by protein, salt, and fluid restriction, in conjunction with hemodialysis. In most cases it is reversible, although sometimes only after protracted periods of oliguria (up to 42 days in one instance in the author's experience). While there is little one can do to reverse the hyperacute rejection associated with the presence of circulating antibodies against the donor's white cells, the more common type of rejection, seen usually—but not exclusively—in the first weeks or months after transplantation is potentially reversible if treated promptly and vigorously. Corticosteroids in large doses (200 to 400 mg. per day of prednisone or its equivalent) either orally or parenterally, are usually effective in aborting or reversing one or more rejection crises, although the dose required to accomplish this may need to be raised in each subsequent episode. Changing the dose of azathioprine for treatment of rejection is impractical, since the full effect of the additional dose may take 48 hours to become manifest. Moreover, the patients are already receiving the amount of drug which is safely tolerated; any significant increase in dose, particularly with declining renal function, may result in potentially fatal bone marrow depression.

If the attempt to reverse the rejection has been successful, renal function should improve and clinical signs disappear soon. Steroid dosage can then be tapered to its previous level over a week or 10 days. Some of the poor results of transplantation in the past have been due to excessive immunosuppressive therapy given to revert rejection beyond the point where this was still possible. If the signs of rejection have not abated in several days, biopsy of the graft is indicated, and if it reveals irreversible morphological changes, the graft should be removed and the patient returned to maintenance dialysis to wait for another transplant.

Besides rejection, the postoperative management of the transplant recipient can be complicated by a variety of problems which are largely the result of the toxic effects of the immunosuppressive agents used. Prominent among these are infections with a large number of pathogens, mainly gram-negative bacteria and opportunistic organisms such as fungi and the herpes and cytomegalovirus.[5, 13, 14] Because the defense mechanisms are significantly impaired by the immunosuppressive agents, these infections are often disseminated and not infrequently fatal. Recently, the development of malignant lymphoma in five patients treated with different immunosuppressive regimens was reported from three widely separated transplantation centers.[32] Although the pathogenesis of these neoplasms is not clear, it is possible that the inhibition of the host's tumor immunity by the immunosuppressive agents used played at least a permissive role in their development.

A curious entity, termed "transplant lung,"[37] which may be due to an autoimmune mechanism, is characterized by alveo-capillary block and hypoxemia, with relatively few clinical findings. It usually accompanies threatened rejection and may be reversed by large doses of steroids. Hypertension may be caused by acute or chronic rejection, steroid medication, salt and water retention, or the recurrence of glomerulonephritis in the graft, an event particularly common in identical twins.

Rarely, it is the result of narrowing of the renal artery at or near the anastomosis site. The appearance of a bruit over this region suggests the diagnosis, which is confirmed by selective arteriography. This lesion is cured by corrective surgery.

Corticosteroids may cause any of their well-known side-effects, including gastrointestinal bleeding, salt and fluid retention, hypertension, psychosis, demineralization of bone, and some less common complications such as aseptic necrosis of the femoral head and acute pancreatitis. Azathioprine often causes marrow depression, manifested chiefly by leukopenia and thrombocytopenia, which is usually—but not always—reversible upon reduction of dose or its temporary discontinuance. In addition, it may also cause liver damage sometimes indistinguishable from hepatitis. Since the probability of marrow depression is greatly increased if the graft is not functioning properly, azathioprine dosage should be reduced in patients with impaired graft function, particularly in the presence of acute tubular necrosis. If the white cell count decreases below 4000, azathioprine should be reduced to 1 to 1.5 mg. per kg. per day, and if below 3000, the patient should be maintained on a very small dose (25 to 50 mg. per day), or the drug should be temporarily discontinued until the bone marrow recovers.

Long-Term Follow-up

If all has gone well, approximately 6 months after transplantation the patient may return home and to the care of his family physician. As mentioned previously, although the probability of acute rejection decreases with each succeeding month after transplantation, occasionally it may occur late, and the physician now in charge of the patient should be alert to this possibility and familiar with its manifestations. A less dramatic but nonetheless serious chronic rejection process may also develop at this stage. It is characterized by a slow decline in kidney function, with development of hypertension accompanied by significant proteinuria. Some of these patients have all the features of the nephrotic syndrome.[12] Morphologically, the kidneys may show marked endothelial vascular proliferation,[33] interstitial cellular infiltrates and fibrosis,[9] or membranous nephropathy.[12] Unfortunately, this late process does not respond to treatment as well as the acute type, and usually progresses to the eventual failure of the graft.

The patient should be seen regularly, perhaps once a month. Although not all the laboratory tests described above are necessary, a urinalysis, complete blood count, and BUN and serum creatinine are mandatory at each office visit. The physician should be encouraged to seek the advice of members of the transplantation team whenever he feels that problems in the patient's management require it. If possible, the patient should continue to be seen once or twice a year at the transplantation center for more elaborate tests. Nonetheless, both his day by day management and the early detection of rejection or drug toxicity are now primarily the responsibility—and the challenge—of his family physician.

PERSPECTIVES

The results of kidney transplantation from all categories of donors have improved consistently over the last years (Fig. 1). If this trend continues, the statistics now being compiled by the Kidney Transplant Registry for its next report will be superior to those discussed in this presentation. Yet, gratifying as it is for those involved in kidney transplantation to see their favorable reports on the "state of the art" outdated while still in the press, it is important to remain aware of how much still remains to be done in kidney transplantation. The average life expectancy of all transplant recipients is not very long, and appears to be even shorter than that of patients on maintenance dialysis. The supply of kidneys, either from living or from cadaver donors, is far short of the demand. Last, graft survival from cadaver donors still leaves much to be desired.

Where can the advances of the next few years be expected to come from? The outcome of kidney grafting from living related donors – which in several centers with much experience approaches 90 per cent survival at 1 and 2 years – is already at a level where further advances are not likely to improve it much. For those, however, not fortunate enough to have a well matched family member willing to donate a kidney, the outlook is far less encouraging. Not only are the results of cadaver kidney transplantation only slightly better than mediocre, but for many, even this chance is not available because of the scarcity of cadaver kidneys. It is therefore in cadaver kidney transplantation that the major efforts of the next few years will have to be concentrated in order to achieve the progress hoped for today. Significant advances may be achieved in several areas.

Organ procurement is inadequate today due to both insufficient public awareness of the benefits which the act of volunteering organs after death may confer on one or more fellow human beings, and to the archaic and complicated legislation which now regulates such donations. In some states permission to use any organ is exclusively the prerogative of the next of kin, even if the deceased has specifically and in writing volunteered his organs before his death. Often, even if permission is granted, during the time required to obtain it the organs deteriorate and can no longer be used. The Uniform Anatomical Gift Act currently under consideration, which, besides making the legal aspects of transplantation uniform, also facilitates donation by making permission by the donor binding after death,[35] should go a long way toward improving the supply of organs. The cooperation of the medical profession, and in particular of those physicians involved in the terminal care of patients who may be adequate donors, will also contribute significantly in this respect.

Organ preservation and tissue typing. Organ banks, where kidneys will be stored as long as needed, seem a rather remote possibility today. On the other hand, the more modest aim of maintaining the kidney viable for 24 hours, while the best recipient can be picked up by histocompatibility testing and the kidney transported to his hospital, seems possible in the very near future.[2] Further improvements in the tech-

niques of tissue typing and increasing the size of the potential recipient pools—by setting up cooperative citywide or even regional typing centers[42]—will greatly enhance the likelihood of superior donor-recipient matches. Should these aims be achieved, the results of cadaver kidney transplantation may, in theory, approach those obtained today with living related donors.

Improved immunosuppression. Antilymphocytic serum promises to be a major addition to the methods of immunosuppression currently available. In contradistinction to all the other agents whose immunosuppressive action is only a by-product of their more toxic effects, antilymphocytic serum is a potent, specific inhibitor of cell-mediated immunological reactions, among them the homograft rejection.[15, 21] However, the translation of the very promising results obtained in laboratory animals to clinical homotransplantation is still not satisfactory, possibly because the preparations now available are not yet adequate. If the technical difficulties of preparing sufficient amounts of potent ALS can be overcome, for example by improving the antigen or by purification of the antibody,[46] significant benefits can be expected from its widespread use in human transplantation in the future.

The last decade has witnessed continuous and very encouraging progress in kidney transplantation. A major breakthrough in any of the areas reviewed above, or, more likely, cumulative slow advances in all of them, will conceivably bring us nearer to what seems utopian today: the day when kidney grafts will be available for all who need them and kidney grafting will confer upon the recipient a near normal life expectancy.

REFERENCES

1. Bach, F. H.: Transplantation: Problems of histocompatibility testing. Science, *159*:1196, 1968.
2. Belzer, F. O., Ashby, B. S., Gulyassy, P. F., and Powell, M.: Successful seventeen-hour preservation and transplantation of human-cadaver kidney. New Eng. J. Med., *278*:608, 1968.
3. Carpenter, C. B., and Austen, K. F.: The early diagnosis of renal allograft rejection. *In* Rapaport, F. T., and Dausset, J., eds.: Human Transplantation. New York, Grune and Stratton, 1968.
4. Carpenter, C. B., Rashid, A., Katz, A. I., Busch, G. J., Sheffer, A. L., Valentine, M. D., Birtch, A. G., Kay, R. E., Austen, K. F., and Merrill, J. P.: Adverse reactions to the use of horse anti-human lymphocyte globulin (ALG). Abstracts, Second Meeting, American Society of Nephrology. Washington, 1968.
5. Craighead, J. E., Hanshaw, J. B., and Carpenter, C. B.: Cytomegalovirus infection after renal transplantation. J.A.M.A., *201*:725, 1967.
6. Elkington, J. R.: Moral problems in the use of borrowed organs, artificial and transplanted (Editorial). Ann. Intern. Med., *60*:309, 1964.
7. Giovanetti, S., and Maggiore, Q.: A low nitrogen diet with proteins of high biological value for severe chronic uremia. Lancet, *1*:1000, 1964.
8. Gleason, R. E., and Murray, J. E.: Report from kidney transplant registry: Analysis of variables in the function of human kidney transplants. II. Immunosuppressive regimens. Transplantation, *5*:360, 1967.
9. Hamburger, J., Crosnier, J., and Dormont, J.: Experience with 45 renal homotransplantations in man. Lancet, *1*:985, 1965.
10. Hamburger, J., and Crosnier, J.: Moral and ethical problems in transplantation. *In* Rapaport, F. T., and Dausset, J., eds.: Human Transplantation, New York, Grune and Stratton, 1968.

11. Hampers, C. L., and Schupak, E.: Long-Term Hemodialysis. New York, Grune and Stratton, 1967.
12. Harlan, W. R., Jr., Holden, K. R., Williams, G. M., and Hume, D. M.: Proteinuria and nephrotic syndrome associated with chronic rejection of kidney transplants. New Eng. J. Med., 277:769, 1967.
13. Hill, R. B., Dahrling, B. E., Starzl, T. E., and Rifkind, D.: Death after transplantation. An analysis of sixty cases. Amer. J. Med., *42*:327, 1967.
14. Hume, D. M.: Renal homotransplantation in man. Ann. Rev. Med., *18*:229, 1967.
15. James, K.: Anti-lymphocyte serum. Clin. Chim. Acta., 22:101, 1968.
16. Kashiwagi, N., Brantigan, C. O., Brettschneider, L., Groth, C. G., and Starzl, T. E.: Clinical reactions and serologic changes after the administration of heterologous antilymphocyte globulin to human recipients of renal homografts. Ann. Intern. Med., *68*:275, 1968.
17. Katz, A. I., Hampers, C. L., and Merrill, J. P.: Secondary hyperparathyroidism and renal osteodystrophy in chronic renal failure. Medicine, *48*:333, 1969.
18. Kincaid-Smith, P., Marshall, V. C., Mathew, T. H., Eremin, J., Brown, R. B., Johnson, N., Lovell, R. R. H., McLeish, D. G., Fairley, K. F., Allcock, E. A., and Ewing, M. R.: Cadaveric renal transplantation. Lancet, 2:59, 1967.
19. Kissmeyer-Nielsen, F., Olsen, S., Petersen, V. P., and Fjeldborg, O.: Hyperacute rejection of kidney allografts, associated with pre-existing humoral antibodies against donor cells. Lancet, 2:662, 1966.
20. Kolff, W. J.: Kidney transplant or home dialysis? Postgrad. Med., *44*:93, 1968.
21. Levey, R. H., and Medawar, P. B.: Some experiments on the action of antilymphoid antisera. Proc. N.Y. Acad. Sci., *129*:164, 1966.
22. Lewis, E. J., Foster, M., DeLa Puente, J., and Scurlock, C.: Survival data for patients undergoing chronic intermittent hemodialysis. Ann. Intern. Med., 70:311, 1969.
23. Maher, J. F., Freeman, R. B., and Schreiner, G. E.: Hemodialysis for chronic renal failure II. Biochemical and clinical aspects. Ann. Intern. Med., *62*:535, 1965.
24. McIntosh, D. A., Peterson, E. W., and McPhaul, J. J., Jr.: Autonomy of parathyroid function after renal homotransplantation. Ann. Intern. Med., *65*:900, 1966.
25. Merrill, J. P.: Letter to the editor. Ann. Intern. Med., *61*:356, 1964.
26. Merrill, J. P.: Medical management of the transplant patient. *In* Rapaport, F. T., and Dausset, J., eds.: Human Transplantation. New York, Grune and Stratton, 1968.
27. Moore, F. D.: Give and Take: The Development of Tissue Transplantation. Philadelphia, W. B. Saunders Co., 1964.
28. Najarian, J. S., Merkel, F. K., Moore, G. E., Good, R. A., and Aust, J. C.: Clinical use of antilymphoblast serum. Transplantation Proceedings, *1*:460, 1969.
29. Patel, R., Mickey, M. R., and Teresaki, P. I.: Serotyping for homotransplantation. Analysis of kidney transplants from unrelated donors. New Eng. J. Med., 279:501, 1968.
30. Patel, R., and Terasaki, P. I.: Significance of the positive crossmatch test in kidney transplantation. New Eng. J. Med., 280:735, 1969.
31. Pendras, J. P., and Erickson, R. V.: Hemodialysis: A successful therapy for chronic uremia. Ann. Intern. Med., *64*:293, 1966.
32. Penn, I., Hammond, W., Brettschneider, L., and Starzl, T. E.: Malignant lymphomas in transplantation patients. Transplantation Proceedings, *1*:106, 1969.
33. Porter, K. A., Thompson, W. B., Owen, K., Kenyon, J. R., Mowbray, J. F., and Peart, W. S.: Obliterative vascular changes in four human kidney homotransplants. Brit. Med. J., 2:639, 1963.
34. Russell, P. S.: Tissue immunity and organ transplantation. Surg. Gynec. Obstet., *126*: 1203, 1968.
35. Sadler, A. M., Sadler, B. L., Stason, E. B., and Stickel, D. L.: Transplantation—A case for consent. New Eng. J. Med., 280:862, 1969.
36. Sixth Report of the Human Kidney Transplant Registry. Transplantation, 6:944, 1968.
37. Slapak, M., Lee, H. M., and Hume, D. M.: Transplant lung—A new syndrome. Brit. Med. J., *1*:80, 1968.
38. Sokol, A., Gral, T., and Rubini, M. E.: Some medical problems of chronic hemodialysis. Calif. Med., *107*:236, 1967.
39. Starzl, T. E., Groth, C. G., Terasaki, P. I., Putnam, C. W., Brettschneider, L., and Marchioro, T. L.: Heterologous antilymphocyte globulin, histocompatibility matching, and human renal homotransplantation. Surg. Gynec. Obstet., *126*:1023, 1968.
40. Starzl, T. E., Marchioro, T. L., Porter, K. A., Faris, T. D., and Carey, T. A.: The role of organ transplantation in pediatrics. Pediat. Clin. N. Amer., *13*:381, 1966.
41. Starzl, T. E., Marchioro, T. L., Porter, K. A., Iwasaki, Y., and Cerilli, G. J.: The use of heterologous antilymphoid agents in canine renal and liver homotransplantation and in human renal homotransplantation. Surg. Gynec. Obstet., *124*:301, 1967.
42. Terasaki, P. I., Mickey, M. R., Siegel, D. P., Mittal, K. K., and Patel, R.: Serotyping for homotransplantation. XX. Selection of recipients for cadaver donor transplants. New Eng. J. Med., 279:1101, 1968.

43. Terasaki, P. I., Vredevoe, D. L., and Mickey, M. R.: Serotyping for homotransplantation. X. Survival of 196 grafted kidneys subsequent to typing. Transplantation (Supp.), 5:1057, 1967.
44. Traeger, J., Carraz, M., Fried, D., Perrin, J., Saubier, E., Bernhardt, J. P., Revillard, J. P., Bonnet, P., Archimbaud, J. P., and Brochier, J.: Studies of anti-lymphocyte globulins made from thoracic duct lymphocytes. Transplantation Proceedings, *1*:455, 1969.
45. Williams, G. M., Hume, D. M., Hudson, R. D., Jr., Morris, P. J., Kano, K., and Milgrom, F.: "Hyperacute" renal-homograft rejection in man. New Eng. J. Med., 279:611, 1968.
46. Woodruff, M. F. A.: Purification of antilymphocytic antibody. Nature, 217:821, 1968.

950 East 59th Street
Chicago, Illinois 60637

New Findings Regarding Autoimmune Factors in Skin Disease

DOROTHY B. WINDHORST, M.D.*

RAYMOND W. STEBLAY, M.D.**

ULRICH H. RUDOFSKY***

JOHN S. THOMPSON, M.D.****

Skin, and more particularly the epidermis, has been regarded by many as a prime target for the development of an immunologic hypersensitivity directed against 'self'.[45] In fact, Whitfield in 1921[41] originally coined the term "autosensitivity" to cover his hypothesis about the nature of certain widespread self-limited eruptions appearing as a late component of severe local tissue damage. Such eruptions, also termed 'id' reactions, are still poorly understood, but certain other skin diseases have recently been the subject of findings which suggest they may be reflecting an attack by immune mechanisms on autologous tissue. It is the purpose of this paper to review these in the context of certain broad advances in the understanding of autoimmune processes affecting another solid tissue, the kidney.

SYSTEMIC LUPUS ERYTHEMATOSUS

Systemic lupus erythematosus (SLE) provides a logical starting point for this discussion, since the renal lesions in this disease are well described and important new information is available regarding the skin

*Assistant Professor, Department of Medicine (Dermatology), University of Chicago
**Assistant Professor and Research Associate, Department of Pathology, University of Chicago
***Research Associate, Department of Pathology, University of Chicago
****Professor of Medicine, University of Iowa College of Medicine, Iowa City, Iowa; formerly Associate Professor, Department of Medicine, University of Chicago

This research was supported by U.S. Public Health Service grants CA-11272-01, AI-00027-10, and HE-04785, and by a grant-in-aid from the Kidney Disease Foundation.

lesions. The kidney disease in LE is regarded as caused by complexes of antigen and antibody which deposit on the glomerular basement membrane (see review by Koffler and Kunkel[23]). This position is supported by immunofluorescent studies which show the presence of gamma globulin and deoxyribonucleic acid (DNA) on the glomerular basement membranes, and by evidence that the gamma globulin which can be eluted from homogenates of affected human kidneys has a high level of specificity for various components of nuclei.[21-24] Additionally, both nuclear antigen and antinuclear antibody can be demonstrated in the serum of patients with SLE.[36] Evidence that such macromolecules can effect damage to the renal glomerulus of experimental animals has been presented by Dixon et al.[13] These investigations also demonstrated that complement components are present in the involved glomeruli of the animal model, as has been repeatedly demonstrated in the kidneys of SLE patients. Complement is generally accepted as a major effector material in the induction of immunologic tissue damage,[25] apparently acting both to damage tissue directly and to attract neutrophilic infiltrates.

Lupus erythematosus, as emphasized by its name, is historically a disease with prominent skin lesions. Dubois and Tuffanelli[14] found that 72 per cent of 520 patients gave a positive history of cutaneous manifestations, and that 10 per cent of the patients had originally been diagnosed as having chronic discoid lupus erythematosus (CDLE). Burnham et al.[5] and Cormane[10] first demonstrated that such skin lesions may exhibit depositions of gamma globulin at the dermal-epidermal junction. Other investigators have confirmed this, and it is now clear that most skin lesions of both discoid and systemic LE contain IgG (Fig. 1*A*). IgM (Fig. 1*B*) and IgA have also been demonstrated, but with less frequency,[12, 31, 39] and the various patterns of fluorescent staining noted have led to several hypotheses regarding evolution of such lesions.[6, 35, 42] Complement is readily demonstrable[20] (Fig. 1*C*), and it is generally accepted that these findings represent evidence for the presence of complement-fixing antibodies at the site of the epidermal basement membrane, analogous to those seen in the renal lesions of LE. The nature of the antigen which is reacting with these antibodies is not established, but two things are worthy of note: no serum reactivity to this area suggestive of anti-basement membrane specificity has ever been noted by indirect fluorescent testing of SLE serum;[42] additionally, Tan and Kunkel[35] demonstrated that one type of globulin-containing aggregate at the dermal-epidermal junction disappeared after treating the tissue sections with DNA-ase, suggesting that the globulin was combined with DNA.

The integration of this information with the hypotheses regarding the renal damage in LE poses some difficulty. If the presence of gamma globulin at the dermal-epidermal junction is a reflection of circulating antigen-antibody complexes, then a filtering or trapping function for such molecules must be suggested for the skin as has been hypothesized for the kidney. This might be acceptable on the basis of purely passive exudation at the site of a pre-existing skin lesion, but the finding of such

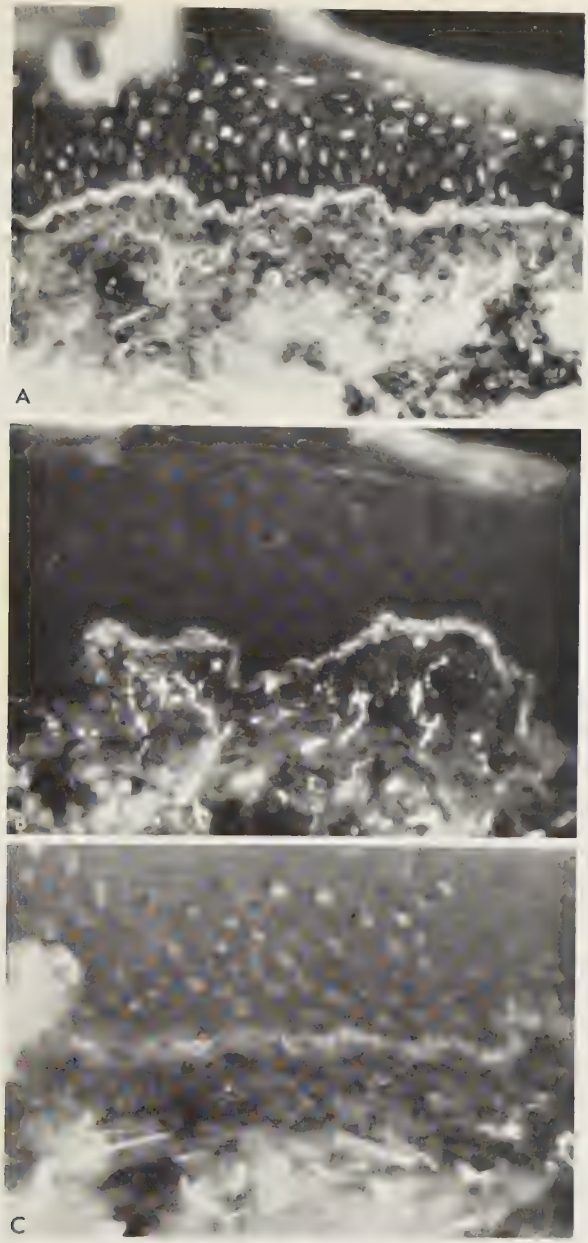


Figure 1. Direct immunofluorescent techniques applied to biopsies of skin lesions of systemic lupus erythematosus (SLE). Chronic discoid LE (CDLE) would look the same. *A*, Fluorescein-labeled anti-human-IgG shows that IgG accumulates in a dermal-epidermal localization. In this patient IgG is also seen on the epidermal nuclei, an infrequent but perhaps important finding (see text). *B*. If another section from the same biopsy is treated with fluorescein labeled anti-human-IgM, it can be seen that IgM is also present at the junction C. Similarly, treatment with fluorescein-labeled anti-human-β_{1c}-β_{1h} shows the presence of complement in the same distribution.

deposits in CDLE, in which, by definition, there are no demonstrable systemic components, must also be explained.

Tuffanelli, Kay, and Fukuyama[42] have shown that 39 per cent of their discoid LE patients had circulating antinuclear factors if the sera were studied using the fluorescent technique with human skin. While such a finding makes it difficult to maintain a clear picture of CDLE as being purely cutaneous, it is certainly provocative in terms of potential understanding of pathogenetic mechanisms.

Tan and Kunkel[35] suggest that skin lesions result from a local phenomenon reflecting formation of antigen-antibody complexes in which the antigen is derived from the nuclear material of the skin, basing their hypothesis on the finding of globulin localized to both dermal and epidermal nuclei in LE skin lesions. In this regard, our findings* agree with those of a number of other authors[6, 15, 30, 31, 39] that the *uninvolved* skin of SLE patients may contain gamma globulin and complement at the dermal-epidermal junction (Fig. 2). All of seven such patients that we have biopsied have this finding, whereas uninvolved skin of discoid LE does not show such deposits, nor did one patient with lupoid hepatitis. Thus, we strongly agree with Percy and Smyth[30] that immunofluorescent study of the normal skin is a valuable diagnostic tool in SLE.

Obviously, the deposition of such material in normal skin, particularly if it could be shown to precede or be correlated with development of new lesions, may have implications as to the pathogenesis of skin lesions. Ten Have-Opbroek[40] has presented evidence that the severity of epidermal change of CDLE lesions correlates well with the prominence of gamma globulin and complement at the dermal-epidermal junction.

Epidermal nuclei do not divide above the basal layer and undergo dissolution at the level of the granular layer. For these reasons, they might logically be regarded as prime targets for antibody directed against nuclear material. In this regard it should be noted that an occasional patient will show, in addition to basement membrane zone globulin deposits, gamma globulin on epidermal nuclei of uninvolved skin as well as in lesions. This does not appear to correlate with the presence of circulating antinuclear factor in the patient biopsied. On the contrary, we have recently observed three such cases in which serum antinuclear factors were not detected by routine procedures but required the use of more sensitive research procedures for their elucidation. Figure 2*A* is taken from one of these patients.

The possibility that epidermal nuclei may serve as a major source of antigen for antibody attack in SLE is certainly an attractive hypothesis, particularly in view of the know proclivity of such patients to react to sun exposure with both skin lesions and systemic exacerbations. Of even greater import, however, is the possibility that altered nuclear material originating in the skin may act to induce the initial formation of such antibodies—that the disease originates in the skin.

Tan and Stoughton[37] demonstrated that exposure of native DNA to ultraviolet light (UV) greatly increased its immunogenicity in rabbits, and that the antibodies which resulted were specific for the UV-DNA

* Unpublished data.

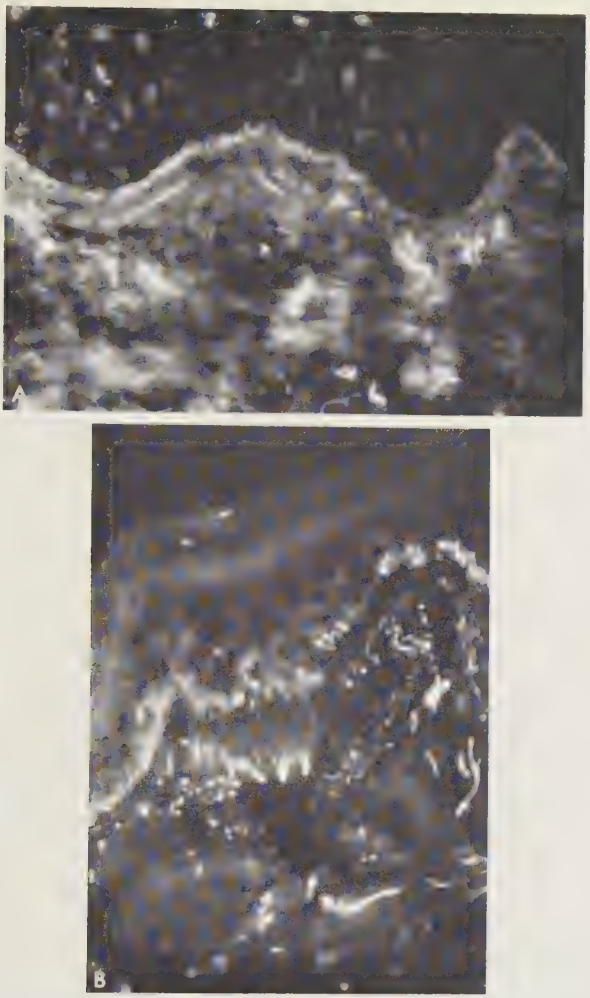


Figure 2. Direct immunofluorescent studies of biopsies of clinically normal skin in SLE. *A*, Taken from a patient who also had skin lesions. *B*, From a patient with fatal renal disease who never had observable cutaneous lesions. Both sections reacted to fluorescein-conjugated anti-human-IgG. The finding of IgG at the dermal-epidermal junction under these conditions appears to have real diagnostic significance (see text). In the original colored pictures, the junction fluoresces bright green, while the dermal fluorescence is nonspecific blue.

over native or heat-denatured DNA. These antibodies, additionally, reacted strongly with the DNA of tissue slices only if the tissues were previously exposed to UV, and the same holds true for the skin of hairless mice irradiated in vivo. The rabbit anti-UV-DNA antibody reacted in a gel-diffusion system only with UV-DNA, and not with native or heat-denatured DNA. Serum from one patient with SLE was shown to react with native DNA and UV-DNA with a line of identity. These same authors have shown that irradiation of human skin renders the nuclei susceptible to combination with the anti UV-DNA antibody.[18]

Thus, it seems possible that at least one target of direct immunologic attack in LE is the nuclear material of the skin. This clearly does not explain the other aspects of this disease process, except, as has been indicated, the renal damage. The kidney appears to be an innocent bystander, suffering damage as a consequence of the deposition of the circulating antigen-antibody complexes and complement fixation. There is no indication that the kidney is the target for antibody attack in the direct sense, even for the antinuclear factors.

GOODPASTURE'S SYNDROME

The best defined situation in which there appears to be a direct immunologic assault against kidney tissue antigens is relatively rare, but is particularly important to discuss in relationship to the recent findings in pemphigus and pemphigoid. True autoimmunity, with a primary attack against kidney tissues, has been demonstrated in Goodpasture's syndrome (and in some cases of glomerulonephritis). Fluorescent microscopy in this severe inflammatory disease of the lungs and kidneys[41] reveals gamma globulin deposition in the basement membranes of the renal glomerulus in a fine linear pattern, in contrast to the granular to lumpy deposition associated with SLE. The alveoli of involved lungs show similar deposits, and the best evidence is that aveolar and glomerular basement membranes share antigens not found in basement membranes of other tissues.[25] The linear staining, seen for both IgG and complement,[22] appears to be directly analogous to the findings in experimental renal disease induced in sheep by immunizing them with glomerular basement membranes isolated from human, monkey, rabbit, and other animal sources.[34] Steblay and Rudofsky[33] have also demonstrated that basement membrane from human lung can induce glomerulonephritis in sheep.

Globulin-containing eluates from kidneys from human Goodpasture's syndrome can induce glomerulonephritis in monkeys. Also, the antibodies directed against basement membrane can be detected in the serum of such patients, particularly if their kidneys have been removed in preparation for homografting. Globulin concentrates of such sera can also induce glomerulonephritis in recipient monkeys.[25]

Thus, the evidence is strong that the affinity of circulating globulins in this process may be indicating a specific immune attack on the tissue involved. This concept is particularly exciting when the new findings regarding two bullous diseases of skin are examined.

PEMPHIGUS VULGARIS AND BULLOUS PEMPHIGOID

Pemphigus vulgaris and bullous pemphigoid are characterized by large flaccid bullae with very little associated inflammatory change. The blisters are very superficial, as indicated by the failure to leave scars on healing and the essentially eroded (rather than ulcerated) nature of the base when the top sloughs. The bullae of pemphigus vulgaris may be

caused to extend laterally by pressure on the contents (Nikolsky's sign), but this is infrequent in bullous pemphigoid. Pemphigus has a greater tendency to involve the mucous membranes with consequent problems in nutrition, but both may be so extensive as to cause the death of the patient secondary to complications of dehydration, malnutrition, and infection. Both are rare in childhood, and pemphigoid is more common in old age. These diseases are threatening to life: they both respond to steroids, but large doses are often necessary.

The essential difference between the two has been a histological one. Bullous pemphigoid is characterized by a separation of the epidermis from the dermis at the site of the basement membrane, whereas typical pemphigus vulgaris exhibits a more complicated picture of acantholysis, or loss of cohesiveness of epidermal cells.[26]

The histology of these diseases has become most relevant in the light of the findings of Beutner et al.[4] regarding pemphigus, and Jordon et al.[18] dealing with pemphigoid. An extensive literature has been admirably reviewed by Beutner, Jordon, and Chorzelski.[3]

The pemphigus story, which has been confirmed by Waldorf et al.[43] and Chorzelski, von Weiss, and Lever,[8] is still unfolding. The serum from patients with this disease can be shown to contain IgG that reacts with squamous epithelium in an intercellular distribution (Fig. 3*A*). Ultrastructural studies by Chorzelski et al.[9] have confirmed the intercellular

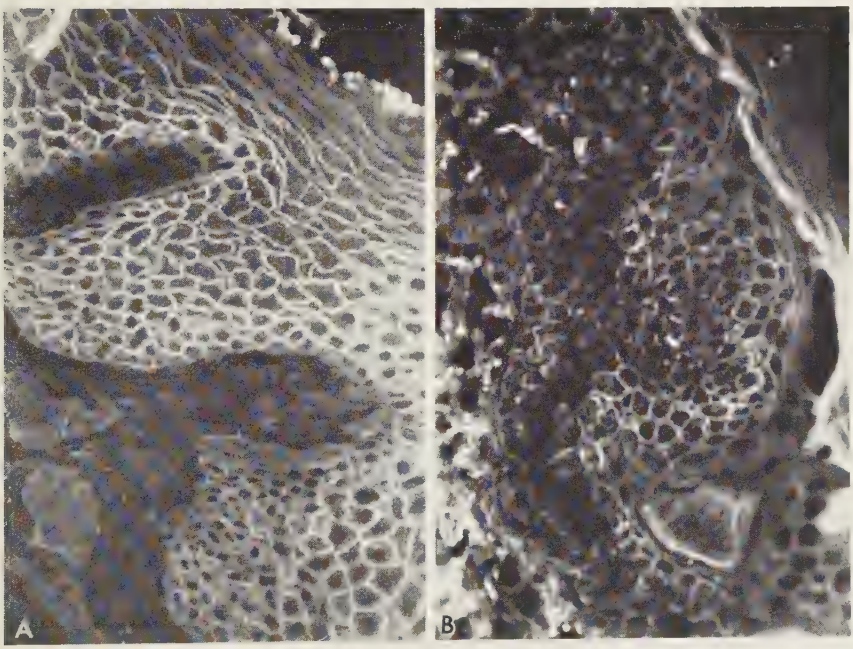


Figure 3. Immunoglobulins in pemphigus vulgaris. *A*, Indirect immunofluorescence study demonstrating that serum from a pemphigus patient contains IgG which has attached to guinea pig esophagus in an intercellular distribution. *B*, Direct study of a biopsy. IgG is seen to be present in a similar distribution on an early pemphigus lesion. (Courtesy of Robert Jordon, M.D., Section of Dermatology, Mayo Clinic.)

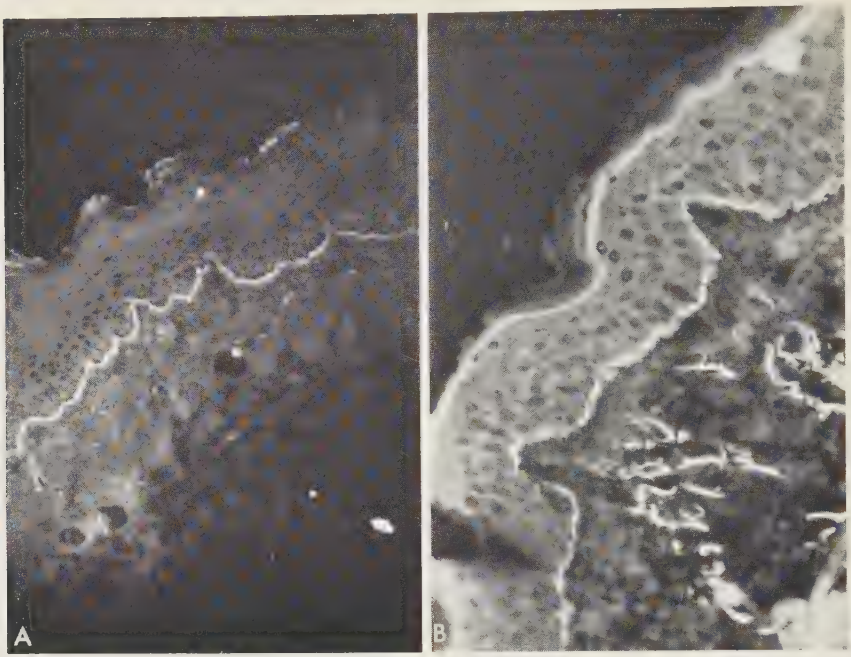


Figure 4. Bullous pemphigoid. *A*, Indirect immunofluorescence indicates that serum from a patient contains IgG which has an affinity for the dermal-epidermal junction of guinea pig esophagus. Direct examination of lesions has shown that IgG is present on the affected skin in a similar distribution. Recently, Jordon et al.[18] have shown that the circulating globulin may be complement-fixing, and that complement may be present on the lesions of pemphigoid. *B*, Direct examination of a lesion for complement. (Courtesy of Robert Jordon, M.D., Section of Dermatology, Mayo Clinic.)

rather than intracellular localization. This reactivity occurs with tissue from various animal sources, such as guinea pig lip and monkey esophagus, suggesting that the material with which the presumed antibody is reacting is tissue specific but not species specific.[1] Direct immunofluorescent microscopy of the lesions reveals a similar distribution of globulin at the edge of blisters (Fig. 3*B*).

The significance of the localization of the gamma globulin in pemphigus, particularly as regards possible pathogenesis, has been enhanced by the findings in bullous pemphigoid. In this condition, the blisters contain IgG deposited at the dermal-epidermal junction, and the serum from patients exhibits this globulin in variable titres (Fig. 4).

Thus, the localization of the gamma globulin in these two diseases follows the histological pattern of the lesion, a feature which has contributed significantly to clarifying our understanding of these two clinically similar diseases.

This approach has proved particularly useful in distinguishing bullous pemphigoid from still other diseases which are sometimes clinically confusing, such as erythema multiforme and bullous dermatitis herpetiformis.[7, 18, 29] Neither of the latter diseases exhibits any tendency to have gamma globulin localized to the lesions in any pattern, nor is there any antibody activity demonstrable in the sera of these patients.

Dermatitis herpetiformis, in particular, has been for many years regarded as closely related to bullous pemphigoid. It is usually a chronic process in which grouped small vesicles are distributed symmetrically over certain areas of predilection. It is characterized by severe pruritus and a specific but poorly understood responsiveness to sulfapyridine and the sulfones. Despite the many distinguishing characteristics of this process, it may sometimes exhibit bullae which are similar to those of bullous pemphigoid, including the histologic finding of a dermal-epidermal separation. These latter features led French dermatologists in particular to consider bullous pemphigoid to be a variant of dermatitis herpetiformis, with bullous dermatitis herpetiformis as an example of a unifying disease.[27] Immunofluorescent studies appear to have resolved the matter in favor of two distinct entities, since dermatitis herpetiformis has not been found to have any of the immunofluorescent characteristics of bullous pemphigoid.

Properly applied, the study of patients' sera by indirect fluorescent microscopy apparently gives few if any false positive results, but false negatives may occur in some 20 per cent of cases. The negative values may not always be false, since it has been demonstrated that the titre of the circulating tissue-reactive globulin detected by the immunofluorescent procedures varies with the severity of the disease process and disappears when the patient is in remission.

The role that these antibodies may play in the pathogenesis of these diseases is of speculative interest. The correlation of titre with severity of disease is reminiscent of the general correlation of disease activity with circulating antinuclear antibody in lupus erythematosus. On the other hand, the situation in Goodpasture's syndrome indicates that severe disease may be associated with very low levels of circulating antibody, presumably because it is preoccupied in a reaction with its tissue antigen. And even in lupus, it seems clear that complement levels correlate better with disease activity than some antibody reactivities.[32]

Complement, indeed, is useful as an indicator of ongoing immune reactions. Its presence on lesions would suggest an element of immunologically mediated tissue damage at the site, and the complement reactivity of the circulating globulins might indicate something as to their pathogenicity. In this regard, Cormane and Chorzelski[11] felt that they could demonstrate bound complement in pemphigus lesions, and Beutner, Jordon, and Chorzelski[3] report early evidence that serum complement levels are low in pemphigus. On the other hand, no demonstration of complement binding capacity of the pemphigus reactive circulating globulins has been achieved, and Pearson and Thompson* found normal complement levels in six pemphigus patients.

In pemphigoid, Jordon[18] has recently demonstrated the capacity of the indirect reacting serum factors to bind complement, and complement can be detected in the direct examination of these lesions (Fig. 4*A*).

Attempts to demonstrate that the circulating antibody may effect lesions on passive transfer have been limited to intradermal injection of

*Unpublished data.

patients' sera into monkey skin. The antibodies fixed to skin in the expected manner, but no lesions were elicited.

Several investigators are indirectly testing the question of autoimmune pathogenesis by immunizing animals with extracts of squamous epithelium, an approach similar to the work using glomerular basement membranes to induce glomerulonephritis. Although this approach is still quite new, it is of interest that two groups have been able to induce antibodies that react similarly to that seen in pemphigus.[2, 3, 16, 17] Inderbitzen and Grob[17] claim to have been able to induce an acantholytic lesion in an immunized animal by the application of freezing trauma.

SUMMARY

Some of the most exciting findings in the field of clinical immunology are unfolding in the area of dermatology. It has now been established that pemphigus and pemphigoid are associated with antibodies having skin specificity. The skin lesions of LE, on the other hand, have been found to contain nonspecific globulin deposits suggestive of the antigen-antibody complex, complement-mediated damage seen in SLE renal disease. The analogy to kidney disease has been drawn as an aid in considering these new data, and serves to illustrate the kind of understanding that must be achieved before an immune pathogenesis of the skin diseases may be acceptable. As a counterpoint, greater insight into the pathogenesis of the skin lesions in lupus erythematosus may be most useful for the additional clarification and amplification of the present hypotheses regarding the systemic manifestations in that syndrome.

REFERENCES

1. Ablin, R. J., and Beutner, E. H.: Absorption studies on antigen(s) of the esophageal mucosa reactive with autoantibodies of pemphigus. Inter. Arch. Allergy Appl. Immunol., *33*:227, 1968.
2. Beutner, E. H., and Ablin, R. J.: Precipitation reactions with epithelial antigens and "pemphigus" heteroantibodies of rabbit and human pemphigus autoantibodies. Fed. Proc., 1968.
3. Beutner, E. H., Jordon, R. E., and Chorzelski, T. P.: The immunopathology of pemphigus and bullous pemphigoid. J. Invest. Derm., *51*:63, 1968.
4. Beutner, E. H., Lever, W. F., Witebsky, E., Jordon, R. E., and Chertock, B.: Autoantibodies in pemphigus vulgaris. J.A.M.A., *192*:682-688, 1965.
5. Burnham, T. K., Neblett, T. R., and Fine, G.: The application of the fluorescent antibody technic to the investigation of lupus erythematosus and various dermatoses. J. Invest. Derm., *41*:451, 1963.
6. Burnham, T. K., and Fine, G.: The immunofluorescent "band" test for lupus erythematosus. Arch. Derm., *99*:413, 1969.
7. Chorzelski, T. P., Jablonska, S., Blaszczyk, M., and Jarzabek, M.: Autoantibodies in pemphigoid. Dermatologica, *136*:325, 1968.
8. Chorzelski, T. P., Von Weiss, J. F., and Lever, W. F.: Clinical significance of autoantibodies in pemphigus. Arch. Derm., *93*:570, 1966.
9. Chorzelski, T. P., Biczysko, W., Dabrowski, J., and Jarzabek, M.: Ultrastructural localization of pemphigus autoantibodies. J. Invest. Derm., *50*:36, 1968.
10. Cormane, R. H.: "Bound" globulin in the skin of patients with chronic discoid LE and systemic LE. Lancet, *1*:534, 1964.
11. Cormane, R. H., and Chorzelski, T. P.: "Bound" complement in the epidermis of patients with pemphigus vulgaris. Dermatologica, *134*:463, 1967.
12. Cormane, R. H.: Classification of immunoglobulins in the dermo-epidermal junction in lupus erythematosus. Clin. Exp. Immun., *1*:207, 1966.
13. Dixon, F. J., Feldman, J. D., and Vazquez, J. J.: Experimental glomerulonephritis. The pathogenesis of a laboratory model resembling the spectrum of human glomerulonephritis. J. Exp. Med., *113*:899, 1961.

14. Dubois, E. L., and Tuffanelli, D. L.: Clinical manifestations of systemic lupus erythematosus. J.A.M.A., *190*:104, 1964.
15. Faille-Kuyper, E. H. B., and Cormane, R. H.: The occurrence of certain serum factors in the dermal-epidermal junction and vessel walls of the skin in lupus erythematosus and other (skin) diseases. Acta Derm. Venereol., *48*:578, 1968.
16. Grob, P. J., and Inderbitzen, T. M.: Experimental production in rabbits of antiepithelial autoantibodies. J. Invest. Derm., *49*:637, 1967.
17. Inderbitzen, T. M., and Grob, P. J.: Destruction of epithelial cells *in vivo* by antiepithelial autoantibodies. J. Invest. Derm., *49*:642, 1967.
18. Jordon, R. E., Beutner, E. H., Witebsky, E., Blumental, G., Hale, W. L., and Lever, W. F.: Basement zone antibodies in bullous pemphigoid. J.A.M.A., *200*:751-756, 1967.
19. Jordon, R. E., Sams, W. M., Jr., and Beutner, E. H.: Complement immunofluorescent staining in bullous pemphigoid. J. Lab. Clin. Med., in press.
20. Kalsbeek, G. L., and Cormane, R. H.: "Bound" complement in the skin of patients with chronic discoid lupus erythematosus and systemic lupus erythematosus. Lancet, 2:178, 1964.
21. Koffler, D., Schur, P. H., and Kunkel, H. G.: Immunological studies concerning the nephritis of systemic lupus erythematosus. J. Exp. Med., *126*:607, 1967.
22. Koffler, D., Sandson, J., Carr, R., and Kunkel, H.: Immunologic studies concerning the pulmonary lesions in Goodpasture's syndrome. Amer. J. Path., *54*:293, 1969.
23. Koffler, D., and Kunkel, H. G.: Mechanisms of renal injury in systemic lupus erythematosus (Editorial). Amer. J. Med., *45*:165, 1968.
24. Krishnan, C., and Kaplan, M. H.: Immunopathologic studies of systemic lupus erythematosus. II. Antinuclear reaction of γ-globulin eluted from homogenates and isolated glomeruli of kidneys from patients with lupus nephritis. J. Clin. Invest., *46*:569, 1967.
25. Lerner, R. A., Glassock, R. J., and Dixon, F. J.: The role of anti-glomerular basement membrane antibody in the pathogenesis of human glomerulonephritis. J. Exp. Med., *126*:989, 1967.
26. Lever, W. F.: Histopathology of the Skin. 3rd ed. Philadelphia, J. B. Lippincott Co., 1961.
27. Lever, W. F.: Pemphigus and Pemphigoid. Springfield, Illinois, Charles C Thomas, 1965.
28. Muller-Eberhard, H. J., Nilsson, U. R., Dalmasso, A. P., Polley, M. J., and Calcott, M. A.: A molecular concept of immune cytolysis. Arch. Path., *82*:205, 1966.
29. Peck, S. M., Osserman, K. E., Weiner, L. B., Lefkovits, A., and Osserman, R. S.: Studies in bullous diseases. New Eng. J. Med., *279*:951, 1968.
30. Percy, J. S., and Smyth, C. J.: Immunofluorescent skin test in systemic lupus. J.A.M.A., *208*:485, 1969.
31. Pohle, E. L., and Tuffanelli, D. L.: Study of cutaneous lupus erythematosus. Arch. Derm., *97*:520, 1968.
32. Schur, P. H., and Sandson, J.: Immunological factors and clinical activity in systemic lupus erythematosus. New Eng. J. Med., *278*:533, 1968.
33. Steblay, R., and Rudofsky, U.: Autoimmune glomerulonephritis induced in sheep by injections of human lung and Freund's adjuvant. Science, *160*:204, 1968.
34. Steblay, R.: Glomerulonephritis induced in sheep by injections of heterologous glomerular basement membrane and Freund's adjuvant. J. Exp. Med., *116*:253, 1962.
35. Tan, E. M., and Kunkel, H. G.: An immunofluorescent study of the skin lesions in systemic LE. Arth. Rheum., *9*:37, 1966.
36. Tan, E. M., Schur, P. H., Carr, R. I., and Kunkel, H. G.: Deoxyribonucleic acid (DNA) and antibodies to DNA in the serum of patients with systemic lupus erythematosus. J. Clin. Invest., *45*:1732, 1966.
37. Tan, E. M., and Stoughton, R. B.: Ultraviolet light alteration of cellular deoxyribonucleic acid *in vivo*. Proc. Nat. Acad. Sci., *62*:708, 1969.
38. Tan, E. M., and Stoughton, R. B.: Ultraviolet light induced damage to desoxyribonucleic acid in human skin. J. Invest. Derm., *52*:537, 1969.
39. Ten Have-Opbroek, A. A. W.: Demonstration of immunoglobulins and complement in the skin of patients with lupus erythematosus. Acta Dermato-Venereol., *46*:68-71, 1966.
40. Ten Have-Opbroek, A. A. W.: On the differential diagnosis between chronic discoid lupus erythematodes and lymphocytic infiltration of the skin. Dermatologica, *132*:109, 1966.
41. Tucker, R. M., and Brown, A. L., Jr.: Goodpasture's syndrome—Clinicopathologic conference. Proc. Mayo Clinic, *43*:449, 1968.
42. Tuffanelli, D. L., Kay, D., and Fukuyama, K.: Dermal-epidermal junction in lupus erythematosus. Arch. Derm., *99*:652, 1969.
43. Waldorf, D. S., Smith, C. W., and Strauss, A. J. L.: Immunofluorescent studies in pemphigus vulgaris. Arch. Dermat., *93*:28, 1966.
44. Whitfield, A.: The question of sensitiveness to nonbacterial toxins and proteins. Brit. J. Derm., *34*:331, 1921.
45. Windhorst, D. B.: Autoimmunity. *In* Criep, L. H., ed.: Dermatologic Allergy. Philadelphia, W. B. Saunders Co., 1967.

950 East 59th Street
Chicago, Illinois 60637

Hemodynamic Observations in Cardiac Transplants

RICHARD A. CARLETON, M.D.*

JAMES G. CLARK, M.D.**

STANLEY J. HELLER, M.D.***

JOHN S. GRAETTINGER, M.D.****

Clinical cardiac transplantation had its dramatic beginning at the hands of Christiaan Barnard on December 3, 1967.[1] Since that time, over 150 efforts to transplant hearts have been undertaken in the world. The preparation for this major therapeutic undertaking had been long, starting with the ingenious work of Carrel and Guthrie in 1905[4] and continuing through the surgical advances of the past two decades.

The purpose of the present communication is to describe some facets of the hemodynamic performance of the transplanted heart. A discussion of cardiac performance following transplantation must be viewed in the over-all context of the problems existing with cardiac transplantation. Essentially, these are five in number.

The first problem which orthotopic homotransplantation posed was that of the sheer magnitude of the surgical undertaking. The ingenious work of Lower, Stoefer, and Shumway provided the world with a means

---

From the Section of Cardio-Respiratory Diseases, Department of Medicine, Presbyterian-St. Luke's Hospital, and the Department of Medicine, University of Illinois College of Medicine, Chicago, Illinois

*Director, Section of Cardio-Respiratory Diseases, Department of Medicine, Presbyterian-St. Luke's Hospital; Professor of Medicine, University of Illinois College of Medicine

**Assistant Attending Physician, Section of Cardio-Respiratory Diseases, Department of Medicine, Presbyterian-St. Luke's Hospital; Instructor in Medicine, University of Illinois College of Medicine

***United States Public Health Service Trainee in Cardiology, Department of Medicine, Presbyterian-St. Luke's Hospital; Research Assistant in Medicine, University of Illinois College of Medicine

****Chairman, Division of Medicine, Presbyterian-St. Luke's Hospital; Professor of Medicine, University of Illinois College of Medicine

Supported in part by a training grant, No. 5T01-HE-05714, from the U.S. Public Health Service.

of making feasible the technique of cardiac excision and subsequent anastomosis in the recipient.[12] This basic surgical technique has subsequently been modified by Cooley and his co-workers to ensure retention of the autorhythmic sinus nodal tissue of the donor heart.[6] The basic surgical technique having been successfully accomplished, residual problems exist in terms of maintaining viability of the donor heart during the transfer from donor to recipient. The present practicable solution to this problem is one of speed during this transfer, to minimize the period of anoxic arrest of the donated heart. The surgical undertaking has proved eminently successful; the immediate surgical mortality of the surgery has been remarkably low. In general, the problem of surgical technology in cardiac transplantation has been solved.

A second major problem in cardiac transplantation is that pertaining to the selection of donors. The entire undertaking of transplantation of unpaired organs has raised a host of moral and ethical factors and has forced society and the medical profession to begin a redefinition of human death. Studies such as those of Rossof and Schwab[16] and essays of the type written by Beecher[2] have greatly clarified the issues and have offered realistic solutions. The problem of donor selection, if not that of donor availability, seems well on the way to solution.

The third area of major question in the realm of cardiac transplantation is that of the choice of recipients. Little information exists in the world pertaining to the prognosis for life of individuals disabled by heart disease. One of the necessary by-products of cardiac transplantation will be the acquisition of meaningful information pertaining to the prognosis of such individuals. Only with such information can one firmly base decisions concerning the applicability and meaningfulness of cardiac transplantation. Preliminary information is now becoming available which suggests that at the present time, minimal if any prolongation of life is afforded by cardiac transplantation when survival of transplant patients is contrasted with that of their peers who, for logistical reasons, did not undergo transplantation.[6] In our own limited experience, of six patients selected for transplantation, one died 3 months after transplantation, one died 6 weeks after selection, and four are alive and clinically improved 10, 9, 6, and 4 months, respectively, after selection. At present, recipient selection remains a process based to a troublesome degree on clinical intuition and acumen. As anterospective data are accumulated concerning the natural history of heart disease, this area in cardiac transplantation will also become more rational.

The fourth problem area pertains to man's presently limited ability to type lymphocytes in particular and tissue in general, the associated limited ability to coordinate activities around the world to ensure the best matching of antigens between potential donor and potential recipient, and his grossly limited ability to counteract the efforts by the body to reject the implanted tissue. The immunologic problems posed by tissue transplantation are immense; the solutions presently are inadequate, but the intensive work proceeding in this field offers hope for continued improvement.

The fifth problem concerns the adequacy of the performance of the transplanted and, of necessity, denervated heart. The hemodynamic

story is one of success for most cardiac transplantations, but as such must be viewed in the context of the four problems mentioned above.

HEMODYNAMIC STATUS OF THE TRANSPLANTED HEART

Lessons from Studies of Denervated Hearts

Two types of studies have provided insight into the capability of hearts to function adequately after transplantation. The first category of observations has been made on the hearts of animals after attempts to denervate the heart. Denervation of the heart has been achieved by a variety of techniques including sympathetic ganglionectomy and vagotomy, complete ablation of mediastinal sympathetic and parasympathetic nerves, and excision and re-implantation (autotransplantation) of the heart. Studies on animals subjected to these procedures have uniformly shown complete denervation after autotransplantation, and variably complete denervation of the heart after division of the autonomic nervous system at various levels. Many of the hemodynamic studies conducted on denervated hearts must be viewed with suspicion unless the heart was totally removed and re-implanted, because the completeness of cardiac denervation is difficult to ascertain with other experimental techniques.[14]

The totally denervated heart has two major physiologic defects. These are loss of control over myocardial contractility and over heart rate. The first is attributable to the striking reduction of the catecholamine content of the myocardium, to approximately 1.2 per cent of normal.[15] Moreover, the uptake of circulating norepinephrine by the denervated heart falls to less than 10 per cent of that of the normal heart.

The first major deficit which exists in the denervated heart is the decrease and loss of variability in myocardial contractility. This variation in contractility is normally induced by sympathetic nervous system stimulation, and constitutes one of the most potent means by which the normal heart adapts to stress. A late response to circulating catecholamines has been shown to occur in the denervated heart under stresses including exercise[3] and systemic hypotension. However, the changes in contractility which have been noted have been small compared to those which occur in the normally innervated heart.

The second major defect in cardiac control concerns heart rate. In the absence of innervation, heart rate remains relatively fixed and somewhat analogous to the intrinsic heart rate described by Jose and Stitt[10] in animals treated simultaneously with a beta adrenergic blocking agent and with atropine. Alteration in heart rate is one of the primary mechanisms for regulating cardiac output; in the denervated or transplanted heart, the utility of this mechanism is minimized. Studies of animals with denervated hearts[14] have shown that heart rate does not respond by acceleration or deceleration even to intense stimulation of the diencephalon. However, the denervated heart remains responsive to circulating catecholamines and will respond after a latent period to diverse types of stimulation in which catecholamines are released in the adrenal gland or at other sympathetic nerve endings.

The nearly fixed heart rate and relatively limited variability in myocardial contractility mean that the denervated heart must rely almost exclusively on the Frank-Starling mechanism to regulate performance. That this somewhat primordial autoregulatory mechanism may be inadequate at times is suggested by the work of Willman, Cooper, and Hanlon,[18] which showed that 17 of 17 dogs who survived more than 2 days after cardiac autotransplantation developed congestive heart failure at some point during the succeeding 3 weeks. Their studies demonstrated markedly reduced catecholamine levels in the myocardium, attesting to the dependence of those hearts on elevated ventricular filling pressures to achieve variation in cardiac output. Chapman and his co-workers,[5] in thorough studies of cardiac performance after total denervation of the canine heart, have shown angiocardiographically that left ventricular end-diastolic volume increases with exercise, which is in contrast to the normal response. This increase in end-diastolic volume corresponds to the increased stroke volume that is the major mechanism by which the denervated heart augments cardiac output during stress.

The deprivation induced by cardiac denervation is not sufficient to be incompatible with life, but does pose limits on the ability of the circulation to respond to demand. Available information demonstrates that cardiac re-innervation occurs, and begins as early as the fourth week after autotransplantation.[14] Information pertaining to re-innervation after homotransplantation is imperfect. The available evidence in dogs suggests that re-innervation may rarely occur but is likely to be incomplete.[11] Studies by Glick et al.[8] after cardiac homotransplantation in man have demonstrated the complete absence of re-innervation 2 months postoperatively.

Hemodynamic Manifestations of Homograft Rejection

The second major consideration in the hemodynamic performance after cardiac homotransplantation pertains to the effects of the rejection process on cardiac performance. Observations by Lower and his co-workers[11] have documented the histologic changes of homograft rejection. Corresponding to the mitochondrial swelling, eventual disruption and necrosis of cell membranes. endothelial damage, and eventual fibrosis, cardiac performance suffers limitations greater than those imposed by denervation alone. One of the early changes noted with cardiac rejection is a decrease in myocardial diastolic compliance.[7] This is manifested by elevated ventricular filling pressure without concomitant increases in end-diastolic volumes. It has been postulated that detection of such changes in diastolic compliance can be used as an early indicator of cardiac homograft rejection in man. Ultrasonic measurements by Harrison and co-workers[9] have shown increased thickness of both the right and left ventricular walls. Their measurements also suggest a decrease in ventricular diastolic compliance related to an inflammatory and infiltrative process in the myocardium as an early phenomenon in the process of rejection. The experiences of others have been similar, but have stressed the more conventional manifestations of congestive failure as indicators of rejection. Thus dyspnea and a decrease in exer-

cise tolerance are common and apparently reliable indicators of established rejection.

Physiologic Data After Human Cardiac Homotransplantation

Physiologic studies after cardiac homotransplantation have been few. Glick, Kinard, and DeBakey[8] have presented studies on three patients; a single such study has been performed in our own laboratory.[3] Several patients have had cardiac output measurements in addition to the routine measurement of the arterial pressure early in their postoperative course. This information has in general shown results which are analogous to those found after canine cardiac homotransplantation. Reports of heart rates have generally been in the range of 90 to 110 per minute. Cardiac output was 4.2, 5.2, and 3.2 liters per minute in the three patients studied by Glick et al.[8] Stinson and his co-workers[17] found normal cardiac outputs in two patients on the first and third postoperative days. Exercise produced slight increases in cardiac output of 1 to 1.3 liters per minute in patients studied by Glick and co-workers. The cardiac output of 8.5 liters per minute at rest rose to 11.5 liters per minute with exercise, as heart rate rose by the fourth minute from 96 to 103 per minute, in the patient studied in this laboratory.

Observations on intracardiac pressures have disclosed that right atrial pressure has been normal to slightly elevated in each patient. The right atrial A wave and P waves have been normal. Tricuspid regurgitation was sought and not found in a single study.[3] Right ventricular pressure at end-diastole has been correspondingly normal to slightly elevated, with detectable increases attributable to A waves. Right ventricular systolic pressure has been variably reported to be elevated or normal. probably reflecting the previous status of the recipient's pulmonary vascular bed. No direct measurements of left atrial pressure have been reported. Pulmonary arterial wedge pressures have been measured in four patients and have been mildly to moderately elevated at rest. In a single observation,[3] the pulmonary arterial wedge pressure rose from a resting mean value of 12 to a level of 18 mm. Hg with exercise. Examination of the pulmonary arterial wedge pressure and of the right atrial pressure in this study disclosed A waves attributable to the contraction of the donated atria.

Left ventricular pressure has been reported as normal at rest in one of the patients reported by Glick et al., and was normal at rest in the patient studied in this laboratory. With elevation of the legs in the latter patient, left ventricular end-diastolic pressure rose from a resting value of 10 to 13 mm. Hg. Supine bicycling exercise further increased end-diastolic pressure to 18 mm. Hg. These changes in left ventricular end-diastolic pressure were reflected in the corresponding increases in left ventricular stroke volume from 47 ml. per beat per square meter at rest to 52 with legs elevated and to 59 with exercise. These observations[3] and those of Glick and others[8] support the concept that the Frank-Starling mechanism is the major means by which the homotransplanted human heart modifies performance.

Left ventricular and pulmonary arterial wedge pressures obtained before and 17 days after cardiac homotransplantation are shown in Figure 1.

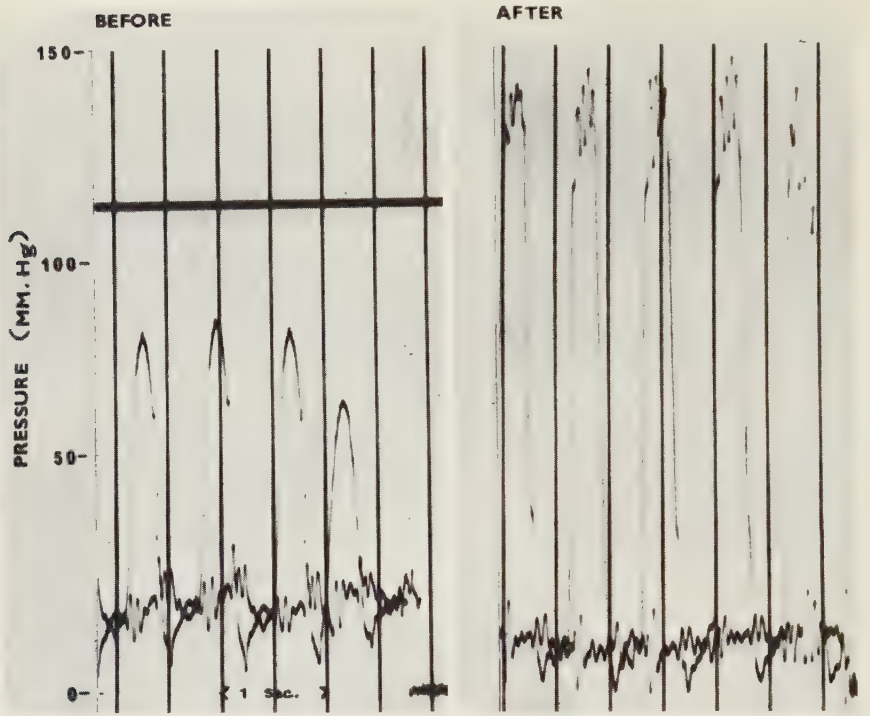


Figure 1. Left ventricular and pulmonary arterial wedge pressures in a patient before and 17 days after cardiac homotransplantation.

The available information pertaining to the performance after cardiac homotransplantation is relatively limited. It suffices, however, to demonstrate that the losses of homeostatic heart rate and of contractility control do not pose insurmountable restrictions. It seems clear that cardiac function of the transplanted heart is adequate to support life and to permit at least moderate degrees of physical activity. The progressive ravages of rejection further limit cardiac capability and, therefore, serve as the limiting factor in the success of human cardiac transplantation.

REFERENCES

1. Barnard, C.: Human cardiac transplantation. An evaluation of the first two operations performed at Groote Schuur Hospital, Cape Town. Amer. J. Cardiol., 22:584, 1968.
2. Beecher, H. K.: Ethical problems created by the hopelessly unconscious patient. New Eng. J. Med., 278:1425, 1968.
3. Carleton, R. A., Heller, S. J., Najafi, H., and Clark, J. G.: Hemodynamic performance of a transplanted human heart. Circulation, in press.
4. Carrel, A., and Guthrie, C. C.: The transplantation of veins and organs. Amer. Med., *10*:1101, 1905.
5. Chapman, C. B., Bruce, T. A., Baker, O., and Fisher, J. N.: The effect of partial and total denervation of the heart on left ventricular function. J. Clin. Invest., 39:976, 1960.
6. Cooley, D. A., Hallman, G. L., Bloodwell, R. D., Nora, J. J., and Leachman, R. D.: Human heart transplantation. Experience with twelve cases. Amer. J. Cardiol., 22:804, 1968.

7. Ecker, R. R., Najem, A. Z., Cyrus, R. J., Webb, W. R., and Sugg, W. L.: Myocardial compliance changes as an indicator of cardiac homograft rejection. Clin. Res., 27:238, 1969.
8. Glick, G., Diethrich, E. B., Kinard, S. A., and DeBakey, M. E.: Physiologic and pharmacologic responses in patients subjected to cardiac transplantation. Clin. Res., 27:242, 1969.
9. Harrison, D. C., Schroeder, J. S., Popp, R., Shumway, N. E., Stinson, E. B., and Dong, E., Jr.: Early diagnosis and treatment of rejection following cardiac transplantation. Clin. Res., 27:244, 1969.
10. Jose, A. D., and Stitt, F.: Cardiac function after combined beta-adrenergic and cholinergic blockade. Relationship of intrinsic rate to contractile force of the heart in dogs. Circulation Res., 20 and 21, Suppl. III, p. 231, 1967.
11. Lower, R. R., Kontos, H. A., Kosek, J. C., Sewell, D. H., and Graham, W. H.: Experiences in heart transplantation. Technic, physiology and rejection. Amer. J. Cardiol., 22:766, 1968.
12. Lower, R. R., Stofer, R. C., and Shumway, N. E.: Homovital transplantation of the heart. J. Thoracic Cardiovas. Surg., *41*:196, 1961.
13. Messmer, B. J., Nora, J. J., Leachman, R. D., and Cooley, D. A.: Survival times after cardiac allografts. Lancet, *1*:954, 1969.
14. Peiss, C. N., Cooper, T., William, V. L., and Randall, W. C.: Circulatory response to electrical and reflex activation of the nervous system after cardiac denervation. Circulation Res., *19*:153, 1966.
15. Potter, L. T., Cooper, T., Willman, V. L., and Wolfe, D. E.: Synthesis, binding, release and metabolism of norepinephrine in normal and transplanted dog hearts. Circulation Res., *16*:468, 1965.
16. Rossof, S. E., and Schwab, R. S.: EEG in establishing brain death. 10-year report, with criteria and legal safeguards in 50 states. Presented at meeting of American Electroencephalographic Society, Atlantic City, June 8, 1967.
17. Stinson, E. B., Dong, E., Jr., Schroeder, J. S., Harrison, D. C., and Shumway, N. E.: Initial clinical experience with heart transplantation. Amer. J. Cardiol., 22:791, 1968.
18. Willman, V. L., Cooper, T., and Hanlon, C. R.: Mechanisms of cardiac failure after excision and re-implantation of the canine heart. Surg. Forum, *13*:93, 1962.

1753 W. Congress Parkway
Chicago, Illinois 60612

III. METABOLISM AND ENDOCRINOLOGY – NEW CONCEPTS AND METHODS

Introduction

If the impact of immunological concepts in medicine needs further emphasis, the following section on metabolism and endocrinology furnishes additional proof either by raising the strong possibility that many metabolic diseases (such as those involving thyroid dysfunction) might be the result of autoimmune reactions or by pointing to newer diagnostic tools developed in this field which are based on immunoassay techniques discussed before. DeGroot's survey on thyroid diseases represents the former; the discussion on hyperparathyroidism and the article on recognition of lipoprotein disturbances, the latter. Dyslipoproteinemias are of course inherited metabolic anomalies. While many are associated with the development of degenerative cardiovascular diseases, the basic genetic metabolic derangement takes precedence over the more popular stress-strain-and-good-living concepts of atherosclerosis and related disorders. A chapter on bone metabolism is included to present a metabolic organ not commonly considered in this light.

H. H. H.

Current Concepts in Management of Thyroid Disease

LESLIE J. DeGROOT, M.D.*

Diagnostic measures used in evaluation of thyroid problems have an air of technical precision about them that should lend security to our clinical judgments. All too often, however, the laboratory reports fail to corroborate our initial impression, and the performance of each additional test adds a new level of sophisticated confusion. We ask the laboratory for an objective assessment of the metabolic level at which the patient's tissues are functioning, and for some clues to the etiology of the disorder. Having been schooled over two decades to believe, and correctly so in most instances, that the basal metabolic rate is an unreliable test of thyroid function, and in the absence of any other clinically useful indicator of tissue metabolic rate, we turn instead to measurement of the level of thyroid hormone in blood, or the level of thyroid gland activity.

The protein-bound iodine is satisfactory in the majority of individuals, but in the very cases in which diagnosis presents a problem, it often fails. For example, in our patient the PBI may be elevated to 11 μg. per 100 ml., but it turns out that she is taking birth control pills. Or we may wish to assess whether the triiodothyronine replacement therapy being given is sufficient, but the PBI measurement is reported to be 1 μg. per 100 ml. and is therefore uninterpretable. The PBI is elevated in patients taking estrogen, because of the increased thyroxine binding globulin capacity in their serum. The PBI is increased to 7 to 11 μg. per 100 ml. in euthyroid persons taking thyroxine replacement therapy because of the strong affinity of the thyroxine for the binding proteins in plasma, and it is low in those on replacement therapy with triiodothyronine because the affinity of this hormone for plasma binding proteins is very weak. Chronic disease, steroid therapy, testosterone treatment, and severe liver disease all may depress the plasma level of thyroid hormone binding proteins

*Professor of Medicine, Department of Medicine, University of Chicago, Chicago, Illinois

Supported in part by U.S. Public Health Service grants Nos. 13377 and 13643, and American Cancer Society grant No. P-298-E.

and thus make the PBI low in the presence of a normal metabolic state. And a large proportion of our patients will have received previous roentgenographic studies with iodinated contrast media, leaving the PBI elevated for weeks or months.

The thyroid radioactive iodine uptake test, while free of the difficulties caused by aberrations of plasma binding proteins, has its own set of limitations. The widespread use of iodate in bread-making has sharply increased the average iodine intake in this country. In many clinics the normal range for the radioactive iodide uptake test at 24 hours has fallen to a level of 14 to 40 per cent. More often than not, further depressions of the RAI uptake are not related to thyroid or pituitary dysfunction, but turn out on careful assessment to be due to ingestion of iodide in medicinal formulations, administration of roentgenographic contrast media, or the use of iodine-containing antiseptics or cosmetics.

While the T_3 resin uptake test, simplified by the introduction of the commercial sponge kits, avoids the pitfalls caused by ingestion of iodide in nonhormonal forms, this test is often difficult to interpret. The T_3 resin uptake test report provides a number which correlates in general with thyroid activity. However, the number can be thought of, in physiologic terms, as the ratio of effective thyroid hormone in blood to the total thyroid hormone binding sites in the blood sample. Since it is a ratio, it must be appreciated that alterations in either of the two factors may cause deviations from the normal range. Thus, in the absence of other data, not even a normal level can be interpreted as an absolute assessment of the metabolic status of the patient. The numerator of this ratio is increased when the supply of free thyroxine in blood is augmented, as in thyrotoxicosis, and is conversely diminished in thyroid failure. The denominator, the level of binding sites in plasma, is made larger by the action of estrogen or occasionally by an inherited genetic trait, and is made smaller by liver disease, severe illness, steroid therapy, testosterone administration, and certain drugs such as diphenylhydantoin (Dilantin).

Even if the tests described above could provide a good answer to our clinical question, often the expense seems excessive to the patient, for such a battery of examinations frequently costs at least $40.

MEASUREMENT OF SERUM THYROXINE

A partial resolution to this diagnostic dilemma is offered by recently introduced techniques that allow estimation of thyroid hormone in the serum (ST_4), even in the presence of iodine contamination, and by calculation of the free thyroxine index (FTI). ST_4 can be conveniently measured by a competitive binding technique, introduced by Ekins (1962) and popularized more recently by Murphy and Pattee (1964). The principle of the test is not difficult. A series of tubes are set up containing (a) an aliquot of a pooled serum which provides a standard quantity of thyroid binding proteins, (b) a standard amount of a resin which will bind free thyroxine, and (c) a standardized amount of radioactive ^{131}I-

labeled thyroxine which serves as a tracer. Standard amounts of unlabeled thyroxine are added to several tubes and, after equilibration, the proportion of the labeled tracer that is displaced from the binding protein onto the resin by the added thyroxine is determined. In such a system uptake of the label by the resin is proportional to the quantity of added thyroxine.

An ethanol extract of the patient's sample is then made, and the extract, which contains thyroxine from the patient's serum, is added to a tube containing the standard pooled serum, labeled thyroxine, and resin. The proportion of the labeled thyroxine displaced by this extract onto the resin is then quantitated and, from the calibration curve, the quantity of unknown thyroxine is estimated. When carefully performed the test is highly reproducible, even in the presence of massive contamination of the serum by iodine in any form.

FREE THYROXINE INDEX

The ST_4 test measures total thyroxine in serum rather than the biologically active hormone, the 0.05 per cent of the total that is not bound to plasma proteins. It is this free hormone which seems to enter cells. The T_3 resin uptake test, described above, represents a ratio of the effective or free serum thyroid hormone to the total binding sites for thyroid hormone in plasma. For calculation of the free thyroxine index, the observed T_3 resin uptake test result is "normalized"—that is, converted to a fraction of the normal level—by dividing the observed value by the average normal control for the laboratory. Multiplication of the patient's serum thyroxine by the patient's normalized T_3 resin uptake produces a number similar in magnitude to the serum thyroxine, which is called the free thyroxine index. It is directly correlated with the level of free thyroid hormone (effective thyroid hormone) in plasma. The clinical utility of the free thyroxine index is attested by a number of publications (Howorth et al., 1969; Arango et al., 1968). It is sufficiently useful to deserve wide clinical application. The FTI is useful because it correlates with the level of free thyroid hormone in blood, and it is obvious that in a wide variety of circumstances the metabolic level is determined not by total ST_4, but by the level of free thyroid hormone.

It should be noted that there are other techniques for arriving at the same sort of answer. Thus the free thyroid hormone fraction can be determined by a column technique (Pileggi et al., 1961) or by the more technically demanding equilibrium dialysis technique (Sterling et al., 1962).

Most laboratories will in time substitute for the routine PBI an ST_4 assay by the competitive binding method. Since materials for the T_3 resin uptake test are available in a convenient commercially supplied kit, it will be easy to combine these two established techniques and achieve a level of diagnostic precision that has been heretofore unavailable. It is also possible that use of this method will allow a reduction in the expense of diagnosis to the patient.

TANNED RED CELL AGGLUTINATION TEST FOR THYROID AUTOANTIBODIES

A third test that should be more widely available is the measurement of antithyroglobulin antibodies by the tanned red cell technique of Boyden (1951). A commercially available kit for measuring these antibodies by a precipitation technique has been available for some time, but is relatively insensitive and of much less clinical utility than the tanned red cell (TRC) agglutination method. Material for performance of the TRC test is also commercially available and is inexpensive (Burroughs Wellcome & Co.). Evidence for the presence of antithyroglobulin antibodies is an important diagnostic point in a number of circumstances. Of patients with Hashimoto's thyroiditis, 60 per cent have a positive agglutination test. If the TRC antibody titer is greater than 1/2560, in a patient who is euthyroid and has diffuse thyroid enlargement, it can be confidently assumed that the histology of the thyroid is that of Hashimoto's thyroiditis. The presence of a positive tanned red cell agglutination test in patients who are euthyroid but have exophthalmos is strong evidence that the patient has Graves' syndrome. Also, positive antibodies are routinely found in patients with primary thyroid failure, but are rare in those whose myxedema is caused by pituitary dysfunction.

CAUSE AND CURE OF GRAVES' DISEASE

Evidence that certain varieties of thyroid disease are associated with, if not caused by, autoimmunity, has accumulated over 50 years. Thymic enlargement and lymph node hyperplasia were recognized shortly after description of Graves' syndrome. In fact thymic resection, combined with subtotal thyroidectomy, was recommended by Crotti (1922) and by Halstead (1914) for therapy of childhood Graves' disease in 1912. Half of their patients were alleged to respond to such treatment. In 1953 elevated levels of gamma globulin were detected in patients with Hashimoto's thyroiditis (Fromm et al., 1953), and subsequently two groups of investigators showed that chronic thyroiditis was associated with serum antibodies directed against antigens present in thyroid tissue (Roitt et al., 1956; Witebsky et al., 1956). Two antigen-antibody systems have been well described and two others are probably present. One form of antibody is directed against thyroglobulin. In a second well described system, circulating antibodies are directed toward a component of the cytoplasmic membranes of the thyroid cell. These antibodies are detected by a complement fixation technique (Trotter et al., 1957), or by Coons' (1950) immunofluorescent method. More recently a factor present in serum of patients with thyrotoxicosis, which causes stimulation of thyroid hormone release, has been implicated as an antibody directed against some component of the thyroid cell. In addition it must be assumed that a cell-mediated autoimmune response occurs in thyroid disease, since the histologic picture in thyroiditis and Graves' disease is so typical of a delayed hypersensitivity type of reaction.

Antibodies detected either by the tanned red cell agglutination or complement fixing techniques occur in about 95 per cent of patients with Hashimoto's thyroiditis and are present in high titers in such individuals. In myxedema the incidence of positive reactions and titers is somewhat lower, and the same applies to patients with Graves' syndrome. Low titers of antibodies also are occasionally found in patients with other thyroid diseases and in a proportion of apparently normal individuals, although even in such persons it is probable that occurrence of antibodies correlates with focal thyroiditis. Antibodies are found in relatives of patients with positive agglutination tests, indicating that the propensity to produce thyroid autoantibodies is inherited as a dominant trait (Hall, 1960). The high titers of the antibodies in patients with primary thyroid failure, the histologic picture of the thyroid in this condition (characterized by fibrosis and lymphocyte infiltration), and the occasional progression of Hashimoto's thyroiditis with goiter to thyroid atrophy, all lead to the supposition that primary thyroid failure represents the end stage of thyroiditis. Further evidence that the antibodies are part of a progressive autoimmune reaction is the positive correlation between the presence of antibodies and the occurrence of myxedema after surgery for Graves' disease (Hjort et al., 1963). Experimental thyroiditis can be produced by immunization of animals with extracts of their own thyroid gland, but the disease can be transferred from one animal to another only by transfer of lymphocytes. Thus a contemporary interpretation is that the tanned red cell agglutinating and complement-fixing antibodies are indicative of an autoimmune reaction, but that a cell-mediated process may be more important in destruction of thyroid cells. The antibodies may be part of a widespread abnormality in immunologic tolerance, since a high proportion of patients with thyroiditis and positive antibody reactions also have serum antibodies directed against gastric mucosal components and may have Addisonian anemia (Doniach et al., 1963). Clinical observations also link thyroiditis with idiopathic adrenal failure, hypophysitis, rheumatoid arthritis, lupus erythematosus, juvenile diabetes, hemolytic anemia, and Sjögren's syndrome.

In 1956, Adams and Purves reported that the plasma of patients with Graves' disease contained a factor which stimulated release of hormone from the thyroid gland of test animals such as the guinea pig. Because of the delayed and prolonged action of this material, it was named long acting thyroid stimulator (LATS). The material is found in 30 to 50 per cent of patients with Graves' disease, and in perhaps 80 per cent if their gamma globulin fraction is concentrated. It occurs rarely in healthy persons. The frequency of positive LATS tests and the levels tend to be higher in patients who have multiple manifestations of Graves' syndrome such as goiter, thyrotoxicosis, exophthalmos, and pretibial myxedema (Lipman et al., 1967). An extensive body of information indicates that the LATS factor is either an immunoglobulin of the G class (7S IgG), or if not, a tiny molecule firmly bound to and inseparable from 7S IgG. LATS can be neutralized by incubation of active sera with antibodies directed against 7S IgG, the gamma chain of IgG, or against the

kappa or lambda light chains. Further, the LATS activity can be absorbed from plasma by incubation with thyroid microsomes, indicating that an immunologic reaction may occur between LATS and a component of the thyroid cell. LATS and the tanned red cell agglutinating or complement-fixing antibodies often coexist in the same individual, but there is little or no correlation between the occurrences of these different antibody systems (Ochi and DeGroot, 1968), indicating that LATS is directed against a separate antigen.

The cause of autoimmunity in Graves' disease is unknown, as is true of all autoimmune phenomena so far recognized. In general the theories can be separated into two categories, those which suggest there is basically a normal immune apparatus in the patient, and those suggesting an abnormality in immune homeostasis. To date no abnormal protein (antigen) that could be indicated as the cause of autoimmunization has been recognized. Further, it appears that the components of the thyroid cell, or at least thyroglobulin, are not antigenically foreign to the host, since 75 per cent of infants and pregnant women have circulating thyroglobulin (Hjort et al., 1962).

Unless some cross-immunity exists between an infectious agent and a thyroid antigen, it is difficult to see how the above data support the concept that the immune apparatus in the patient with thyroid autoimmunity is basically normal, but is stimulated by a foreign or abnormal antigen. Rather, the genetic factor in inheritance of thyroid autoimmunity, the coincident occurrence of other autoantibodies, and association with a variety of other presumed autoimmune diseases, indicate that the defect is an intrinsic abnormality of immune homeostasis. There is also discussion as to whether the antibodies are purely secondary markers following upon cell destruction, are primary factors in the disease, or simply mediate a part of the process. As of this reading, it seems necessary to grant that autoantibodies at least mediate the cell destruction of Hashimoto's thyroiditis and the thyroid hyperplasia and hypersecretion of Graves' disease. Perhaps the antibodies occur because of some genetically determined abnormality of immune homeostasis, or are conditioned by environmental factors. A scheme relating these agents to the autoimmune diseases of the thyroid is given in Figure 1.

LATS is known to cause release of thyroid hormone from the animal or human gland. In addition it stimulates every phase of thyroid tissue activity, including the uptake of radioactive iodide, oxidation of glucose, formation of phospholipids, synthesis of nucleic acids and protein, and pinocytosis of colloid (McKenzie, 1968). If given to animals over several days, it can result in formation of goiter, thus proving it is a growth stimulant to the gland (Ochi and DeGroot).

It is by no means as clear that LATS has any relation to exophthalmos. Many patients have severe exophthalmos without measurable LATS in their serum. Administration of LATS has not been shown to produce experimental exophthalmos. Localization of LATS in orbital tissues (DeGroot and Ochi) or absorption of LATS from plasma by orbital tissue fragments (Shillinglaw, 1968) does not occur. In fact, it must be admitted, with considerable disappointment, that while LATS has pro-

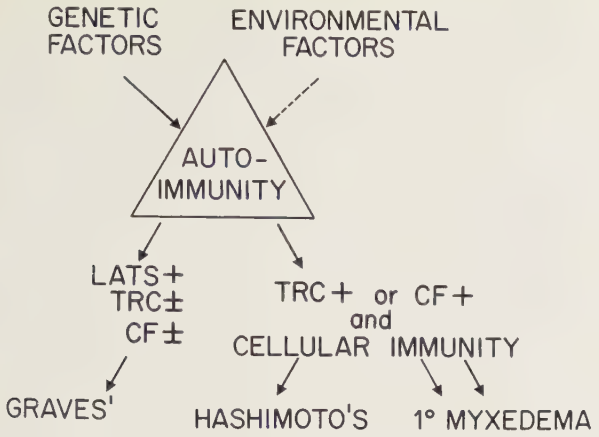


Figure 1. Scheme of the relations of auto-antibodies and other factors in thyroid diseases.

vided a fascinating new chapter in understanding the Graves' disease phenomenon, it has had little or no impact on treatment of the illness. Bioassay of LATS does play a useful role in diagnosis in some instances. Thus, a positive LATS assay in an individual with unilateral exophthalmos constitutes strong evidence that the disease process is an expression of the Graves' syndrome rather than a locally invasive process of the orbit. Although a decrease in LATS activity correlates in a general way with control of the Graves' process, regardless of the form of therapy, the occurrence of or titer of LATS has not been found useful in choosing a specific form of treatment or as a guide to therapy.

CHOICE OF TREATMENT IN GRAVES' DISEASE

As in most medical problems whose etiology and pathophysiology are not completely understood, choice of appropriate therapy remains a topic of spirited and even emotional debate. We are fortunate in having at least three excellent modalities of treatment, including prolonged administration of antithyroid drugs, or the more "definitive" therapies such as radioactive iodide administration or subtotal thyroidectomy. It must be admitted that each of these treatments is directed at a symptom of the Graves' syndrome rather than at the basic etiologic mechanism. Whether we administer drugs or radioactive iodine, or take out a portion of the thyroid, in each instance the therapy is designed to temporarily suppress secretion of thyroid hormone into plasma. If the therapy is successful, the derangement driving the thyroid in uncontrolled secretory activity seems to abate, and in months or years the gland is able to re-establish a normal homeostatic relationship with the pituitary and to supply the body with physiologic amounts of thyroid hormone.

The chief limitations of antithyroid drug therapy are the prolonged period of supervision necessary (although this may be changing) and the uncertainty of final success. Studies done in the 1950's and 1960's indicated that antithyroid drug treatment of 1 or 2 years' duration was correlated with a higher incidence of prolonged remission than was a briefer course of treatment. Thus it has been customary in most clinics to treat for 1 or 2 years. During this time the patient must be frequently monitored for the occurrence of adverse reactions such as fever, pruritus, hives, neutropenia, or agranulocytosis.

Since there is a general aversion to surgery in children unless it is clearly necessary, and because there is strong evidence that radiation to the child's thyroid may induce malignancy, antithyroid drug treatment remains the therapy of choice in children and in young adults. Usually therapy is initiated with 10 mg. of methimazole or 150 mg. of propylthiouracil every 8 hours, and this dose is increased if it is not effective within 1 to 3 weeks. Rarely, doses up to 40 or 80 mg. of methimazole daily are required. Ideally an improvement in the clinical status is apparent within 2 to 3 weeks, and euthyroidism is achieved within 4 to 6 weeks. At this point the dosage is progressively reduced to a level which maintains the ST_4 in the normal range, or else desiccated thyroid is added to provide normal hormone supply while hormone production by the gland is blocked by the antithyroid drug. After 1 or 2 years of control, the medication is gradually discontinued. Perhaps 50 per cent of patients enter a prolonged remission.

Greer and Smith (1954), and Werner et al. (1955) observed that the elevated radioactive iodide uptake of the thyroid in Graves' disease was not suppressed by administration of exogenous thyroid hormone, in contrast to the response of the normal gland. Astwood (1962) and Cassidy (1965) recognized that patients who went into remission after antithyroid therapy had a decrease in size of the goiter and a return of normal thyroid suppressibility. More recently Harden and co-workers (Alexander, 1967) noted that in some patients thyroid suppressibility returned shortly after institution of antithyroid drug therapy, often within 6 to 10 weeks. It is possible that such patients may need a relatively brief course of antithyroid therapy, or in fact that they may be the ones who will respond to medical therapy with a prolonged remission. Studies are being conducted in numerous clinics at this time to determine whether the return of thyroid suppressibility can be a useful predictor of patients who should receive antithyroid drugs rather than some other form of treatment.

Antithyroid drug therapy is also useful in pregnancy. It seems advisable in such patients to use the lowest dose of antithyroid drug that will keep the patient's metabolic status at a "high normal" level, with a PBI from 9 to 11 μg. per 100 ml., so that there will be less possibility of making the fetus hypothyroid.

In some instances antithyroid drugs are useful as a prelude to treatment with radioactive iodine—for example, if there is fear that RAI might induce the release of hormone from the gland and exacerbate

coincident heart disease. Likewise, antithyroid drug therapy is routinely used prior to surgery.

RAI Therapy

After its introduction in the 1940's, the ease and simplicity of radioactive iodine therapy shortly made it the most commonly employed modality of treatment. As experience accumulated, it became clear that radioactive iodine therapy was not associated with an increased incidence of leukemia (Saenger et al., 1968), or in significant gonadal exposure to radiation. There is, however, a lingering concern that RAI could lead to development of thyroid tumors long after the radiation therapy, and this fear has restricted its use in most clinics to adults. Since thyroid neoplasia as a result of radiation exposure is very rare in adults, in contrast to the recognized association in children (Hempelman, 1968), and because thyroid growth and its accompanied mitotic activity is probably ended in humans by age 18, it seems entirely safe to administer radioactive iodine to patients above age 25. It is recognized that this age limit is entirely arbitrary, and it will be moved upward or downward as experience dictates.

Unfortunately a high proportion of patients who have received radioactive iodine appear to develop permanent hypothyroidism. In fact, with the doses of radioactive iodine employed up to a few years ago, 10 or 20 per cent were hypothyroid within 1 year and 2 to 3 per cent became hypothyroid annually thereafter (Dunn et al., 1964). This high incidence of thyroid failure has been attributed to derangement of the normal reproductive capacity of the cells, leading to progressive atrophy of the gland, or to blood vessel damage leading to fibrosis. In an effort to avoid this complication, some clinics now employ doses of radioactive iodide significantly lower than initially used, in the order of 50 microcuries of ^{131}I retained per gram of thyroid at 24 hours. This dose of radiation is insufficient in many patients to control thyrotoxicosis promptly, and therefore RAI treatment must be supplemented by the use of another agent such as methimazole (or propylthiouracil), or administration of iodine (Hagen et al., 1967). After a period of months, the ancillary treatment may be withdrawn and the patients then commonly remain euthyroid. It is probable that the lower-dose treatment will increase the necessity for second or third doses of radiation, although this has not been borne out by data so far available. Another hazard is that in many patients the thyrotoxicosis will be allowed to linger for a prolonged period before control is achieved. This can be avoided by the judicious administration of antithyroid drugs or potassium iodide as a supplement to radiation, and in patients who, because of severe depletion, coincident heart disease, or for other reasons, require prompt relief from their illness, it is legitimate to use larger quantities of radioactive iodide and thus insure prompt control of thyrotoxicosis (Hamburger, 1964). In such patients it may be anticipated that the larger dose of radioactive iodide will be associated with a higher incidence of permanent hypothyroidism, but this can readily be controlled by replacement therapy and is a small price to pay for prompt control of thyrotoxicosis.

Surgical Therapy

Subtotal resection of the thyroid provides control of the disease in perhaps 95 per cent of patients and is entirely safe when the subject is rendered euthyroid by pretreatment with antithyroid drugs, and if vascularity of the gland is suppressed by administration of potassium iodide, and possibly thyroid hormone. About 20 per cent of patients ultimately become hypothyroid, but the incidence of side-effects such as hypoparathyroidism and vocal cord paralysis is gratifyingly low when the operation is done by skilled surgeons. We must note that the decline in frequency of operations has brought with it a coincident decrease in the supply of skilled surgeons. Surgical therapy is appropriate for patients under age 25 who have failed to respond to an adequate trial with antithyroid drug therapy, who react adversely to the drugs, or who are irresponsible in their treatment. It is successfully used for control of thyrotoxicosis during the second trimester of pregnancy. The only real contraindication is previous surgery, which makes a second operation difficult, often bloody, and increases the frequency of complications. About the most damaging thing that can be said against surgery is that it is usually an operation not required.

OPHTHALMOLOGIC COMPLICATIONS

There appears to be no correlation between the type of therapy offered for control of thyrotoxicosis and the subsequent course of the eye complications, at least if the patients are not rendered hypothyroid (Hamilton, 1967; Aranow et al., 1965). Most individuals with Graves' syndrome have the worst manifestations of their ocular disease (chemosis, proptosis, edema, ocular paresis, and decreased visual acuity) at the time of recognition of their illness, and the eyes tend to improve with therapy. In a small number of patients, perhaps 2 or 3 per cent of the total, progressive worsening of exophthalmos occurs despite adequate control of the hyperthyroid state. In these patients there is a baffling array of possible therapeutic approaches, none of which can be said with certainty to guarantee relief. Because the usual natural history is one of progressive improvement and ultimate resolution, it is best to seek time by application of local measures, such as protection of the cornea with dark glasses and wide-brimmed hats, sleeping with the head elevated, use of diuretics, placing Saran Wrap or compresses over the eyes at night, methylcellulose drops, and antibiotic ointment for conjunctivitis. Most patients, even with severe exophthalmos, weather the storm adequately with such measures.

For those who develop a progressive decrease in vision, something more dramatic must be done. Most thyroidologists would next use a course of treatment with prednisone, commonly 40 mg. daily, for a period of 4 to 8 weeks. Some patients respond with remission in their symptoms, up to and including complete return to normalcy. Others, for reasons unknown, will have no response. A few of these may be benefited by the use of massive doses, such as 100 or 200 mg. of prednisone daily, but

obviously this carries with it the attendant risk of iatrogenic Cushing's syndrome, gastric ulceration, diabetes, etc. If steroids are ineffective and vision is progressively impaired, operative decompression by the Krönlein technique is probably the best next procedure. Lateral tarsorrhaphy may be useful to protect the cornea in some patients, but should not be thought of as offering protection against damage to the optic nerve or retina from severe proptosis. Word-of-mouth reports indicate that some patients have been treated with azathioprine (Imuran), 150 to 300 mg. daily, for prolonged periods, but that this therapy has not been associated with impressive resolution of exophthalmos.

Catz (1965) has written extensively about the use of thyroid ablation to influence the course of exophthalmos (Bauer et al., 1966). According to his hypothesis, the thyroid provides antigenic stimulus for the production of LATS or other factors, which are the mediators of the infiltrative ophthalmopathy of Graves' disease. Destruction of the thyroid by combined use of total thyroidectomy, followed by radioactive iodide treatment under TSH stimulation, is reported to control severe progressive exophthalmos. In this very difficult management problem, even minor improvement can be judged a clinical triumph. And because of the vagaries of the illness and the natural tendency for resolution, a large carefully controlled series is necessary to prove the effectiveness of any therapeutic modality. The few publications which have appeared from other clinics reporting attempts to assess the role of thyroid ablation in the control of exophthalmos do not indicate general usefulness for this therapy (Volpe et al., 1969; Werner et al., 1967). In fact, some individuals seem to be made worse by thyroid ablation (Pinchera et al., 1969), and it is suggested that this is due to the release of thyroid antigens into the circulation after radioiodine administration. Certainly LATS titers do tend to be increased transiently after radioactive iodide destruction of thyroid tissue (Kriss et al., 1967). For the moment it is difficult to take a strong position in favor of or against thyroid ablation for control of exophthalmos. Rather, it will be necessary for greater clinical experience before this therapy can be properly evaluated.

USE OF OTHER DRUGS TO CONTROL SYMPTOMS

Some patients present dramatic exacerbations of thyrotoxicosis, often in association with intercurrent illness or an operative procedure, to an extent that merits the name thyroid crisis. In such patients it would obviously be useful to have an agent which acutely suppresses the action of thyroid hormone in the peripheral tissues, rather than to depend on a suppression of production of hormone through the use of antithyroid drugs or radioactive iodine. Iodine is useful since it promptly suppresses release of thyroid hormone into blood, and will, in patients with Graves' disease, depress the PBI within a day or two. In patients who have fairly severe thyrotoxicosis and in whom both prolonged and prompt control is desirable, it is often useful to give radioactive iodine and then within 10 to 14 days to add potassium iodide.

If there is concern about late exacerbation of symptoms by the potassium iodide—and this phenomena certainly does occur in long term treatment of Graves' disease patients with iodine—the patient may first be treated with RAI and at 10 to 14 days be given antithyroid drugs and then potassium iodide. The suppression of thyroid hormone release by KI works perfectly well in the presence of a methimazole block, apparently because it is dependent upon a high level of iodide ion within the thyroid cell rather than upon the production or binding of large amounts of organic iodine in the gland (Greer et al., 1956). In patients who present with severe thyroid crises, it is best to institute therapy with an antithyroid drug in high doses, for example methimazole 20 mg. or propylthiouracil 200 mg. every 6 hours, and then to add Lugol's solution—two or more drops twice daily, beginning within 1 or 2 hours after administration of the antithyroid drug. A diminution in hypermetabolism may be apparent within 1 to 3 days with such combined therapy.

More recently there has been interest in measures designed to control the apparent heightened sympathetic activity characteristic of the thyrotoxic state. Of course phenobarbital has been administered for years and is effective in controlling hyperkinetic behavior in some patients. Reserpine and guanethidine have also been used. These agents probably act by depleting tissue stores of catecholamines, and reserpine has as well a central sedative effect. Doses of reserpine (Canary et al., 1957) such as 1 to 3 mg. intramuscularly every 8 hours or of guanethidine (Howitt et al., 1966) 20 to 50 mg. three times daily are reported to be useful in the management of severe thyrotoxicosis. Phenoxybenzamine, an alpha sympathetic blocker, offers little therapeutic utility when used alone.

A rash of reports indicate that propranolol, a beta adrenergic blocker, is more specific in the control of symptoms of thyrotoxicosis (Weiner et al., 1969; Wilson et al., 1964, 1966). There is no doubt that propranolol, in doses of 20 to 80 mg. four times daily, can cause a reduction in tachycardia due to sinus rhythm or to atrial fibrillation. It also appears to slow down the hyperdynamic circulation of thyrotoxicosis and induces a prolongation of the ejection time of the ventricle. It is not as clear that propanolol lowers the cardiac output or decreases cardiac work. Recent reports indicate that patients have an improved feeling of well being during propranolol therapy (Shanks et al., 1969), and that combined therapy with propranolol and phenoxybenzamine lowers the BMR, causes weight gain, and reverses many clinical features of thyrotoxicosis (Stout et al., 1969). It is clear that propranolol is useful in controlling tachycardia if this is imperative in the management of severe thyrotoxicosis. It may have more far reaching effects on the circulatory system, and in fact on the general metabolism of the hyperthyroid individual. The drug (2 mg. intravenously or 20 to 80 mg. by mouth) caused rapid alleviation of fever, tachycardia, and even coma in a few patients with thyroid storm (Das et al., 1969) and interestingly was noted to rapidly reduce elevated free fatty acids to normal in two desperately ill patients (Buckle, 1968).

It will require considerable study before the exact role of β adrenergic blockade in management of thyrotoxicosis is defined. On the basis

of present data, this measure is useful for relief of symptoms and control of tachycardia. The dramatic effects reported in thyroid storm justify its use in this often fatal illness. Since β adrenergic blockade can precipitate congestive heart failure, this possible adverse action must be carefully monitored during therapy.

REFERENCES

Alexander, W. D., Harden, R. M., Shimmins, J., McLarty, D., and McGill, P.: Development of thyroidal suppression by triiodothyronine during six months treatment of thyrotoxicosis with antithyroid drugs. J. Clin. Endocrinol., 27:1682, 1967.

Arango, G., Mayberry, W. E., Hockert, T. J., and Elveback, L. R.: Total and free human serum thyroxine in normal and abnormal thyroid states. Mayo Clin. Proc., *43*:503, 1968.

Aranow, H., and Day, R. M.: Management of thyrotoxicosis in patients with ophthalmopathy: Antithyroid regimen determined primarily by ocular manifestations. J. Clin. Endocrinol., 25:1, 1965.

Astwood, E. B.: The Thyroid. New York, Harper & Row, 1962, p. 684.

Bauer, F., and Catz, B.: Radioactive iodine therapy for progressive malignant exophthalmos. Acta Endocrinol., *51*:15, 1966.

Boyden, S. V.: Adsorption of proteins on erythrocytes treated with tannic acid and subsequent hemagglutination by antiprotein sera. J. Exper. Med., *93*:107, 1951.

Buckle, R. M.: Treatment of thyroid crisis by beta-adrenergic blockade. Acta Endocrinol., 57:168, 1968.

Canary, J. J., Schoaf, M., Duffy, B. J., and Kyle, L. H.: Effects of oral and intramuscular administration of reserpine in thyrotoxicosis. New Eng. J. Med., 257:435, 1957.

Cassidy, C. E.: Use of a thyroid suppression test as a guide to prognosis of hyperthyroidism treated with antithyroid drugs. J. Clin. Endocrinol., 25:155, 1965.

Catz, B., and Perzik, S. L.: Subtotal vs. total ablation of thyroid: Malignant exophthalmos and its relation to remnant thyroid. *In* Cassano, C., and Andreoli, M., eds.: Current Topics in Thyroid Research. New York, Academic Press, 1965.

Coons, A. H., and Kaplan, M. H.: Localization of antigen in tissue cells. J. Exper. Med., *91*:1, 1950.

Crotti, A.: Thyroid and Thymus. Philadelphia, Lea & Febiger, 1922.

Das, G., and Krieger, M.: Treatment of thyrotoxic storm with intravenous administration of propranolol. Ann. Int. Med., 70:985, 1969.

Doniach, D., Roitt, I. M., and Taylor, K. B.: Autoimmune phenomena in pernicious anemia. Brit. Med. J., *1*:1374, 1963.

Dunn, J. T., and Chapman, E. M.: Rising incidence of hypothyroidism after radioactive-iodine therapy in thyrotoxicosis. New Eng. J. Med., *271*:1037, 1964.

Ekins, R.: Saturation analysis: A microanalytical technique for assaying some compounds of biological importance. Radioaktive Isotope in Klinik und Forschung, Band V (Vortrage am Gasteiner Internationalen Symposion 1962), 1963, pp. 211-220.

Fromm, G. A., Lascano, E. F., Bur, G. E., and Escalente, D.: Tiroiditis cronica inespecifica. Rev. Asoc. Med. Arg., 67:162, 1953.

Greer, M. A., and DeGroot, L. J.: The effect of stable iodide on thyroid secretion in man. Metabolism, 5:682, 1956.

Greer, M. A., and Smith, G. E.: Method for increasing the accuracy of the radioiodine uptake as a test for thyroid function by the use of desiccated thyroid. J. Clin. Endocrinol., *14*:1374, 1954.

Hagen, G. A., Ouellette, R. P., and Chapman, E. M.: Comparison of high and low dosage levels of ^{131}I in the treatment of thyrotoxicosis. New Eng. J. Med., 277:559, 1967.

Hall, R., Owen, S. G., and Smart, G. A.: Evidence for genetic predisposition to formation of thyroid autoantibodies. Lancet, 2:187, 1960.

Halsted, W. S.: The significance of the thymus gland in Graves' disease. Bull. Johns Hopkins Hosp., 25:223, 1914.

Hamburger, J. I., Kadian, G., Rossin, H. W., and Sills, R.: I^{131} therapy for hyperthyroidism: A plea for less conservatism in complicated cases. J. Nuclear Med., 5:675, 1964.

Hamilton, R. D., Mayberry, W. E., McConahey, W. M., and Hanson, K. C.: Ophthalmopathy of Graves' disease: A comparison between patients treated surgically and patients treated with radioiodide. Mayo Clin. Proc., *42*:812, 1967.

Hempelman, L. H.: Risk of thyroid neoplasm after irradiation in childhood. Science, *160*:159, 1968.

Hjort, T., Jeppesen, F., Okholm, K., and Temler, J.: The prognostic significance of thyroid autoantibodies in the development of postoperative myxedema. Danish Medical Bulletin, *10*:159, 1963.

Hjort, T., and Pedersen, G. T.: Thyroglobulin antibodies and thyroglobulin in the serum in pregnant and parturient women and in newborn infants. Lancet, *2*:259, 1962.
Howitt, G., and Rowlands, D. J.: Beta sympathetic blockade in hyperthyroidism. Lancet, *1*:628, 1966.
Howorth, P. J., and MacLagan, N. F.: Clinical application of serum-total-thyroxine estimation, resin uptake, and free-thyroxine index. Lancet, *1*:224, 1969.
Kriss, J. P., Pleshakov, V., Rosenblum, A. L., Holderness, M., Sharp, G., and Utiger, R.: Studies on the pathogenesis of the ophthalmopathy of Graves' disease. J. Clin. Endocrinol., *27*:582, 1967.
Lipman, L. M., Green, D. E., Snyder, N. J., Nelson, J. C., and Solomon, D. H.: Relationship of long-acting thyroid stimulator to the clinical features and course of Graves' disease. Amer. J. Med., *43*:486, 1967.
McKenzie, J. M.: Humoral factors in the pathogenesis of Graves' disease. Physiological Reviews, *48*:252, 1968.
Murphy, B. P., and Pattee, C. J.: Determination of thyroxine utilizing the property of protein-binding. J. Clin. Endocrinol., *24*:187, 1964.
Ochi, Y., and DeGroot, L. J.: Studies on the immunological properties of LATS. Endocrinology, *83*:845, 1968.
Ochi, Y., and DeGroot, L. J.: Stimulation of thyroid hyperplasia and protein synthesis by LATS. Endocrinology, *85*:344, 1969.
Pileggi, V. J., Lee, N. D., Golub, O. J., and Henry, R. J.: Determination of iodine compounds in serum. I. Serum thyroxine in the presence of some iodine contaminants. J. Clin. Endocrinol., *21*:1272, 1961.
Pinchera, A., Liberti, P., Martino, E., Fenzi, G. F., Grasso, L., Rovis, L., and Baschieri, L.: Effects of antithyroid therapy on the long acting thyroid stimulator and the anti-thyroglobulin antibodies. J. Clin. Endocrinol., *29*:231, 1969.
Purves, H. D., and Adams, D. D.: An abnormal thyroid stimulator in the sera of hyperthyroid patients. *In* Pitt-Rivers, R., ed.: Advances in Thyroid Research (Trans Intl. Goitre Conf., London, July 1960). Oxford, Pergamon Press, 1961, pp. 184-188.
Roitt, I. M., Doniach, D., Campbell, P. N., and Hudson, R. V.: Auto-antibodies in Hashimoto's disease (lymphadenoid goitre): Preliminary communication. Lancet, *2*:820, 1956.
Saenger, E., Thoma, G., and Tompkins, E.: Incidence of leukemia following treatment of hyperthyroidism. J.A.M.A., *205*:855, 1968.
Shanks, R. G., Lowe, D. C., Hadden, D. R., McDevitt, D. G., and Montgomery, D. A.: Controlled trial of propranolol in thyrotoxicosis. Lancet, *1*:933, 1969.
Shillinglaw, J., and Utiger, R. D.: Failure of retro-orbital tissue to neutralize the biological activity of the long-acting thyroid stimulator. J. Clin Endocrinol., *28*:1069, 1968.
Sterling, K., and Hegedus, A.: Measurement of free thyroxine concentration in human serum. J. Clin. Invest., *41*:1031, 1962.
Stout, B. D., Wiener, L., and Cox, J. W.: Combined alpha and beta sympathetic blockade in hyperthyroidism. Ann. Int. Med., *70*:963, 1969.
Trotter, W. R., Belvayin, G., and Waddams, A.: Precipitating and complement fixing antibodies in Hashimoto's disease. Proc. Roy. Soc. Med., *50*:961, 1957.
Volpe, R., Desbarats-Schonbaum, M. L., Schonbaum, E., Row, V. V., and Ezrin, C.: The effect of radioablation of the thyroid gland in Graves' disease with high levels of long-acting thyroid stimulator (LATS). Amer. J. Med., *46*:217, 1969.
Werner, S. C., Feind, C. R., and Aida, M.: Graves' disease and total thyroidectomy. New Eng. J. Med., *276*:132, 1967
Werner, S. C., Spooner, M., and Hamilton, H.: Further evidence that hyperthyroidism (Graves' disease) is not hyperpituitarism: Effects of triiodothyronine and sodium iodide. J. Clin. Endocrinol., *15*:715, 1955.
Wiener, L., Stout, B. D., and Cox, J. W.: Influence of beta sympathetic blockade (propranolol) on the hemodynamics of hyperthyroidism. Amer. J. Med., *46*:227, 1969.
Wilson, W. R., Theilen, E. O., and Fletcher, F. W.: Pharmacodynamic effects of beta-adrenergic receptor blockade in patients with hyperthyroidism. J. Clin. Invest., *43*:1697, 1964.
Wilson, W. R., Theilen, E. O., Hege, J. H., and Valenca, M. R.: Effects of beta-adrenergic receptor blockade in normal subjects before, during and after triiodothyronine-induced hypermetabolism. J. Clin. Invest., *45*:1159, 1966.
Witebsky, E., Rose, N. R., and Shulman, S.: Studies of normal and malignant tissue antigens. Cancer Res., *16*:831, 1956.

950 East 59th Street
Chicago, Illinois 60637

Primary Hyperparathyroidism: A Simplified Approach to Diagnosis

ERIC REISS, M.D.*

In general, the larger the number of tests proposed for the diagnosis of a disease, the lower the specificity of each. By this criterion, the nonspecificity of tests for primary hyperparathyroidism is impressive. Despite the confusing number of diagnostic procedures proposed, clinical practice among experienced endocrinologists is remarkably uniform. The diagnosis of primary hyperparathyroidism can be established with high probability by simple clinical evaluation.

NATURAL HISTORY

In some instances, hyperparathyroidism causes unequivocal evidence of end-organ damage (renal calculi, nephrocalcinosis, and osteitis fibrosa). This may be called symptomatic hyperparathyroidism. At other times, the disease presents as hypercalcemia without evidence of end-organ damage (asymptomatic hyperparathyroidism). Recognition of these different forms is important for therapy.

Symptomatic hyperparathyroidism requires prompt surgical management since there is good evidence that cure of the hormone excess decreases calculus formation and permits healing of bones.[2, 12] By contrast, appropriate therapy for asymptomatic patients cannot be recommended with certainty since there exists no reliable information about the natural history of this form of the disease. Patients with established hyperparathyroidism often give histories of having passed urinary calculi 20 or 30 years earlier. Was the disease present for all this time without causing overt harm, or was the initial episode of stone formation unrelated to parathyroid hyperfunction? In many instances, hypercalcemia has been documented to be present for many years without any recognizable clinical manifestations. On the other end of the spectrum, there are patients in whom hyperparathyroidism appears to be associated with rapidly progressive calcareous renal disease or osteitis fibrosa.

*Chairman, Department of Medicine, Michael Reese Hospital and Medical Center; Professor of Medicine, Department of Medicine, University of Chicago, Chicago, Illinois

In view of these uncertainties, most clinicians recommend curative surgery for asymptomatic as well as symptomatic hyperparathyroidism. However, if cooperation in long-term follow-up observations can be assured, there is probably no urgency about surgery in asymptomatic patients with minimal hypercalcemia. Surgery may then be performed if hypercalcemia becomes more severe or organ damage becomes evident. Severe hypercalcemia (13 mg. per 100 ml. or above) deserves more prompt attention since the rate of development and severity of end-organ damage are probably related to the degree of hypercalcemia.

DIAGNOSIS

A strong case can be made for concentrating diagnostic efforts on patients in whom prompt diagnosis and treatment have a high probability of real benefit. In practice, this means patients with calcareous renal disease, nephrocalcinosis, osteitis fibrosa, or symptoms related to hypercalcemia. These are the patients demanding prompt therapy.

However, apart from these specific stigmata of damage, the clinical manifestations of the disease are protean.[7, 11] Acid-peptic disease, pancreatitis, hypertension, crystal deposit disease in joints, affective disorders, and a host of other less common disorders have been associated with primary hyperparathyroidism. It is not clear that treatment of the associated hyperparathyroidism will modify the course of these diseases.

With increasing use of multichannel biochemical screening, diagnostic studies for hyperparathyroidism are becoming commonplace in patients manifesting no evidence of the disease. Hyperparathyroidism discovered in this manner provides an unusual opportunity for long-term prospective studies of the natural history of patients with asymptomatic hypercalcemia.

Acid-peptic disease as a concomitant of hyperparathyroidism deserves special attention. Peptic ulceration occurs in patients with hyperparathyroidism only slightly more frequently than in the general population.[14] However, hyperparathyroidism is common in peptic ulceration when both conditions occur as part of a syndrome of familial endocrine adenomatosis or the Zollinger-Ellison syndrome. For this reason, the incidence of acid-peptic disease in large series of patients with hyperparathyroidism appears to depend on the extent to which the series contains patients having these syndromes. Patients with peptic ulcer and hyperparathyroidism deserve further study, including careful evaluation of gastric secretory patterns and roentgenographic examination of the sella turcica.

Serum Calcium

For practical purposes, hypercalcemia is the sine qua non of diagnosis. Normocalcemic hyperparathyroidism has been recorded on several instances, but one must be careful about defining this condition. How does one classify the disease if two of a dozen calcium determinations show slight hypercalcemia while 10 are normal? Normocalcemic hyper-

parathyroidism, defined by consistent normocalcemia over many determinations, is rare indeed.

All experienced physicians place prime diagnostic emphasis on carefully measured serum calcium concentrations. It is essential that the physician be able to rely on the laboratory. The absolute range of normal is less critical than knowing what the range is in a particular laboratory. Absolute values depend on the method and the reference standard used. In some laboratories 10.5 mg. per 100 ml. is the upper limit of normal; in others it is 11.0 mg.

In the experience of most, hypercalcemia is often intermittent rather than constant in hyperparathyroidism. However, Dent has pointed out that, when appropriate precautions are taken, hypercalcemia can be consistently demonstrated: "One quite normal value for total calcium obtained with impeccable technique makes the diagnosis most unlikely."[9] To obtain such consistent data, results must be normalized for specific gravity of the serum. Since this is not common practice, fluctuations of serum calcium from normocalcemia to hypercalcemia are frequent in ordinary practice.

Differential Diagnosis of Hypercalcemia

When hypercalcemia is discovered accidentally in an asymptomatic patient, the diagnosis of hyperparathyroidism is made by exclusion. In Table 1, the principal diseases associated with hypercalcemia are listed. The only condition that is readily mistaken for hyperparathyroidism is an occult malignancy. Multiple myeloma, sarcoidosis, the milk-alkali syndrome, vitamin D intoxication, and hyperthyroidism should be detected by thorough clinical evaluation. On rare occasions, hyperparathyroidism coexists with another disease that can produce hypercalcemia. Under these circumstances, hypercalcemia will persist despite effective therapy of that disease. For example, hypercalcemia of sarcoidosis is promptly relieved by corticosteroid therapy. If hypercalcemia persists despite therapy, coexistent hyperparathyroidism should be suspected.

Hypercalcemia occurs commonly in the course of various malignancies. In some, the rate of bone destruction by metastases is believed to be so great that the kidneys' capacity to excrete the excess calcium is exceeded. Since the renal excretion capacity for calcium is high, this simple explanation is probably only partially correct. A negative radiographic survey of bones does not provide assurance that skeletal metas-

Table 1. *Differential Diagnosis of Hypercalcemia*

Primary hyperparathyroidism
Cancer with and without metastases
Multiple myeloma
Sarcoidosis
Milk-alkali syndrome
Hyperthyroidism
Immobilization in children and patients with Paget's disease

tases do not exist, since the skeleton may be extensively replaced by tumor without providing a radiographic clue.

In many malignant diseases with hypercalcemia, skeletal metastases are undetectable, even after careful postmortem examination. In some of these patients, the tumor elaborates a substance or substances that are physiologically and immunologically akin to PTH.[19, 22] Bronchogenic carcinoma and hypernephroma are the tumors that are most commonly complicated by hypercalcemia in the absence of skeletal metastases, but many other tumors are known on occasion to give rise to this type of hypercalcemia.

In practice, the differentiation between asymptomatic hyperparathyroidism and occult malignancy complicated by hypercalcemia is exceedingly difficult, and at times impossible. In many malignant conditions, hypercalcemia is relieved by hydrocortisone therapy (50 mg. every 8 hours for 10 days).[23] In primary hyperparathyroidism, this regimen does not affect the hypercalcemia.[10] This is a valuable test. If a thorough clinical evaluation, including history, physical examination, routine laboratory data, chest films, complete roentgenographic examination of the gastrointestinal tract, and intravenous pyelography are normal and steroids fail to decrease the serum calcium, primary hyperparathyroidism can be assumed to be present. Other special metabolic tests, including radioimmunoassay of serum PTH, provide no help in this differential diagnosis.

Serum Phosphorus

Hypophosphatemia was once considered an important diagnostic clue. Increasing experience, however, has resulted in recognition that the serum phosphorus has only limited usefulness in differential diagnosis. Although hypophosphatemia is common in hyperparathyroidism, it is also often absent. Moreover, it also occurs in occult malignant conditions with hypercalcemia and thus provides no assistance in the most difficult diagnostic dilemma. With one exception, the clinical conditions associated with hypophosphatemia, listed in Table 2, are not likely to be confused with primary hyperparathyroidism. The exception, *idiopathic hypercalciuria*, is characterized by normocalcemia and considerably more pronounced hypercalciuria than occurs in primary hyperparathyroidism.

Table 2. *Conditions Associated With Hypophosphatemia*

Primary hyperparathyroidism
Long term dietary deficiency; alcoholism; therapy with phosphate-binding gels
Malabsorption syndromes
Vitamin D-resistant rickets
Fanconi syndrome
Idiopathic hypercalciuria

Skeletal Studies

Subperiosteal resorption of bone in the phalanges and resorption of the distal ends of the clavicles are characteristic. When these changes

are observed in a hypercalcemic patient, they are virtually diagnostic of hyperparathyroidism. When the bone disease is severe, the serum alkaline phosphatase is nearly always elevated. Although advanced bone disease in primary hyperparathyroidism is rare in today's practice, its detection is important since surgical cure of the hormone excess results in healing of bone with dramatic alleviation of skeletal symptoms.

Urinary Calcium

Measurement of the urinary calcium is of limited value in the differential diagnosis of hyperparathyroidism. PTH increases the tubular reabsorption of calcium.[24] Hence, for any particular concentration of serum calcium, the hyperparathyroid subject should excrete less calcium in the urine than the euparathyroid subject. This prediction from physiologic studies has been documented in human disease.[15] Therefore, hypercalciuria in hyperparathyroidism generally means only that hypercalcemia is marked.

The measurement of urinary calcium is of value, however, in the patient with renal calculi who has equivocal hypercalcemia. In such a patient, marked hypercalciuria favors a diagnosis of idiopathic hypercalciuria. Since the urine calcium is relatively insensitive to dietary calcium intake, the common custom of collecting specimens for measurement after several days on a low calcium diet wastes hospital days. Acid should always be added to the collection jars to prevent precipitation of calcium phosphate before the specimen reaches the laboratory.

Specialized Metabolic Tests

While it is not the purpose of this presentation to review in detail the many specialized metabolic tests that have been proposed,[21] a few remarks may serve to clarify the issue. These tests have merit in the experience of those who take a research interest in them. It often happens that others who apply these tests only sporadically obtain less satisfactory results. The general problem may be illustrated with reference to the tubular reabsorption of phosphate (TRP).

If the test is properly performed, the TRP is nearly always low in patients with primary hyperparathyroidism.[6] But the restriction "properly performed" is a serious one. In normal man, the TRP is sensitive to dietary intake of phosphorus. When the dietary intake decreases, the TRP increases; when the dietary intake increases, the TRP decreases. Thus, phosphate homeostasis is maintained by the exquisite adjustments of renal mechanisms. The normal TRP must be defined in relation to a particular level of phosphate intake. If the phosphate intake is unknown, no normal value can be assigned. This simple and obvious restriction in interpreting the TRP has vitiated many studies.

Urinary phosphate excretion is so delicately attuned to intake that phosphate homeostasis is normally maintained with only small changes of serum phosphorus concentration despite marked fluctuations of dietary intake. Failure of the serum phosphate to remain within normal limits reflects either unusually severe and prolonged restriction of intake or failure of the homeostatic mechanisms. For this reason, it has been suggested that the normal TRP should be interpreted not only in terms

of intake, but also in terms of prevailing serum phosphorus concentration.[13] A TRP of 80 per cent, for example, may be normal when the serum phosphorus concentration is 3.5 mg. per 100 ml. but abnormally low in a patient with a serum phosphorus concentration of 1.5 mg. per 100 ml. Thus, as in the case of urinary calcium, the normal pattern of urinary phosphate excretion must be interpreted in relation to the prevailing serum phosphorus concentration.

This analysis makes it apparent that a random TRP, without reference to dietary intake and prevailing serum phosphorus concentration, represents a meaningless datum. It is not surprising, therefore, that physicians report normal TRP in the presence of established primary hyperparathyroidism. They have simply failed to understand the rationale of the test and the factors that normally control phosphate excretion.

One wonders how often such simple considerations have lessened the diagnostic value of various tests. The originator of the test is nearly always aware of pitfalls; the imitator is not.

RADIOIMMUNOASSAY OF PTH

The technique of radioimmunoassay as developed by Berson and Yalow has revolutionized clinical endocrinology.[3] Although application of this powerful tool to the study of PTH in human material has proved difficult and some conflicting results have been reported, consistent data have recently been obtained.[3, 4, 16, 18] With the use of a sufficiently sensitive antiserum, a clear separation can be obtained between normal and hyperparathyroid sera. There is hope that the assay will establish the diagnosis of hyperparathyroidism with assurance in patients with borderline hypercalcemia. However, it is already evident that high circulating concentrations of PTH occur under conditions other than primary hyperparathyroidism and that assay results must be carefully correlated with other clinical and laboratory data. Two examples are given.

Increased quantities of immunoassayable PTH have been measured in patients with certain malignancies.[19] We have recently noted that approximately two thirds of the sera from hypercalcemic patients with malignant disease contain increased quantities of material that immunologically reacts like PTH (unpublished data). This observation seriously limits the usefulness of the assay in diagnosis.

The sensitivity of circulating PTH to diminished renal function has recently been documented.[17] Small decreases of the glomerular filtration rate are associated with unequivocal elevations of PTH. Although the mechanisms by which PTH secretion is increased in early renal disease are not known, a plausible hypothesis has been formulated.[5] When nephrons are lost in the course of progressive renal disease, maintenance of normophosphatemia requires that the TRP decrease. Phosphate retention causes a small increase in serum phosphorus and, reciprocally, a small decrease of serum calcium. This stimulates PTH secretion, which lowers the TRP and restores normocalcemia and normophosphatemia.

Homeostasis appears to be maintained at the expense of some degree of hyperparathyroidism.

At the present, the immunoassay of PTH is of greatest interest as a tool in clinical investigation. For example, we have measured serum PTH in patients with primary hyperparathyroidism before and during induction of severe hypercalcemia by an intravenous calcium infusion.[18] In patients who harbored a single parathyroid adenoma, serum PTH was not suppressed by calcium infusion. By contrast, in patients with primary hyperplasia of two or more parathyroid glands, serum PTH was suppressed by calcium infusion. These observations suggest that an adenoma functions as a true neoplasm in that it is not susceptible to feedback control over a wide range of elevated serum calcium concentrations. However, primary hyperplasia appeared to be characterized by maintenance of the normal control system. The physiologic abnormality in hyperplasia appears to be an abnormally high setting of the control system. There is no information concerning factors influencing the level at which the parathyroid-calcium control operates.

TREATMENT

When the decision for curative surgery has been made, the selection of a surgeon becomes critical. Many physicians lacking extensive experience with primary hyperparathyroidism believe that any surgeon who is expert in neck surgery can perform the task adequately. This is a misconception. Adequate parathyroid surgery requires extensive experience, thorough understanding of the pathologic anatomy of the parathyroid glands, and the assistance of experienced parathyroid pathologists.

If the disease is extensive and obvious, with marked hypercalcemia, and is caused by a solitary large parathyroid adenoma, surgery is likely to be simple and accomplished with ease. However, in many instances, the pathology encountered at the operating table may be difficult to interpret, and the proper surgical procedure is often difficult. In addition, recognition of abnormal parathyroid tissue requires experience. Aberrantly located parathyroid glands may be difficult to find and require the services of a surgeon experienced in the search for them.

Examples of poorly performed surgery are all too common. The problem of appropriate parathyroid surgery is not nearly as simple today as it was formerly. The most trying problem for the surgeon is the entity of parathyroid hyperplasia giving rise to primary hyperparathyroidism. Hyperplasia involving predominantly the water-clear cells was first described in 1934;[1] that involving primarily the chief cells escaped detection until 1958.[8] It has recently been recognized that hyperplasia is a common pathologic entity giving rise to primary hyperparathyroidism.[20] Parathyroid hyperplasia may be difficult to recognize at the operating table. Not all glands are involved by the hyperplastic process to the same extent. Some may be markedly enlarged; others only slightly. The pathologist is unable to differentiate between hyperplasia and adenoma on examination of a single gland. This difficulty is apparent not only

at the time of frozen section but also when the permanent sections are examined.

Lack of uniformity of pathologic interpretation of parathyroid sections has complicated the field. At present, the pattern of experienced surgical practice is changing. In our experience, it is best to aim for locating all four parathyroid glands. Identification of glands should be accomplished by biopsy and frozen section. Small fat pads and lymph nodes are readily mistaken for parathyroid glands, even by experienced parathyroid surgeons. There is no substitute for the biopsy.

The complexities of modern day parathyroid surgery permit the general surgeon to refer patients with primary hyperparathyroidism to colleagues who take a special interest in the surgical treatment of this disease. There is no stigma attached to such a referral. In fact, the surgeon making such a referral deserves our respect and admiration.

Postoperative Care

In the majority of patients, postoperative management is simple. If a thorough parathyroid exploration has been performed, some degree of tetany is expected. There occurs some suppression of parathyroid function, which is characterized by mild hypocalcemia and hyperphosphatemia. This is self-limited if appropriate surgery has been performed. The remaining parathyroid tissue will regain its function after a few weeks. Vigorous treatment with calcium and vitamin D should be avoided. Most patients can tolerate minimal hypocalcemia symptoms, such as numbness and tingling in the hands and about the mouth. Hypocalcemia stimulates the parathyroid remnant to regain normal function. Excessively vigorous treatment of hypocalcemia symptoms delays the return of normal parathyroid function.

If the patient has advanced bone disease, prolonged and severe tetany is the rule. Hypocalcemia is associated with hypophosphatemia since both minerals are rapidly taken up by the depleted bones. This condition has been called "recalcification tetany" or the "hungry bone syndrome." Vigorous treatment with oral calcium, large doses of vitamin D, and oral phosphate is required to alleviate the symptoms of tetany. At times, prolonged therapy with intravenous calcium may be required to alleviate the disabling symptoms.

SUMMARY

The diagnosis of primary hyperparathyroidism is accomplished by the correlation of clinical findings with simple laboratory determinations. Hypercalcemia is a hallmark of the disease, and diagnosis involves predominantly the differentiation between hyperparathyroidism and other diseases that can give rise to hypercalcemia.

In patients with renal calculi or skeletal complications of hyperparathyroidism, surgical cure of the disease in unequivocally indicated. The natural history of the disease in asymptomatic patients is unknown. According to current practices, surgical treatment is indicated if hypercalcemia is severe. Patients with minimal or equivocal hypercalcemia may be followed carefully on a long-term basis.

Special metabolic testing is of limited usefulness in the differential diagnosis of hypercalcemia. A radioimmunoassay for circulating PTH has provided interesting new information in clinical investigation, but is not required in the routine diagnosis.

Surgical treatment of primary hyperparathyroidism requires experience and understanding. It is best accomplished by a team of physician, surgeon, and pathologist who have taken a special interest in the problem.

REFERENCES

1. Albright, F., Bloomberg, E., Castleman, B., and Churchill, E. D.: Hyperparathyroidism due to diffuse hyperplasia of all the parathyroid glands rather than adenoma of one: Clinical studies on 3 such cases. Arch. Int. Med., *54*:315, 1934.
2. Albright, F., and Reifenstein, E. C., Jr.: The Parathyroid Glands and Metabolic Bone Disease. Selected Studies. Baltimore, Williams & Wilkins Company, 1948.
3. Berson, S. A., and Yalow, R. S.: Radioimmunoassays of peptide hormones in plasma. New Eng. J. Med., 277:640, 1967.
4. Berson, S. A., and Yalow, R. S.: Parathyroid hormone in plasma in adenomatous hyperparathyroidism, uremia and bronchogenic carcinoma. Science, *154*:907, 1966.
5. Bricker, N. S., Slatopolsky, E., Reiss, E., and Avioli, L. V.: Calcium, phosphorus, and bone in renal disease and transplantation. Arch. Int. Med., *123*:543, 1969.
6. Chambers, E. L., Jr., Gordon, G. S., Goldman, L., and Reifenstein, E. C., Jr.: Tests for hyperparathyroidism: Tubular reabsorption of phosphate, phosphate deprivation, and calcium infusion. J. Clin. Endocrinol., *16*:1507, 1956.
7. Cope, O.: Hyperparathyroidism: Diagnosis and management. Amer. J. Surg., *99*:394, 1960.
8. Cope, O., Keynes, W. M., Roth, S. I., and Castleman, B.: Primary chief-cell hyperplasia of the parathyroid glands: A new entity in the surgery of hyperparathyroidism. Ann. Surg., *148*:375, 1958.
9. Dent, C. E.: Some problems of hyperparathyroidism. Brit. Med. J., *2*:1419, 1962.
10. Dent, C. E., and Watson, L.: The hydrocortisone test in primary and tertiary hyperparathyroidism. Lancet, 2:662, 1968.
11. Keating, F. R.: Diagnosis of primary hyperparathyroidism. J.A.M.A., *178*:547, 1961.
12. McGeown, M. G.: Effect of parathyroidectomy on the incidence of renal calculi. Lancet, *1*:586, 1961.
13. Nordin, B. E. C., and Fraser, R.: Assessment of urinary phosphate excretion. Lancet, *1*:947, 1960.
14. Ostrow, J. D., Blanshard, G., and Gray, S. J.: Peptic ulcer in primary hyperparathyroidism. Amer. J. Med., *29*:769, 1960.
15. Peacock, M., Robertson, W. G., and Nordin, B. E. C.: Relation between serum and urinary calcium with particular reference to parathyroid activity. Lancet, *1*:384, 1969.
16. Reiss, E., and Canterbury, J. M.: A radioimmunoassay for parathyroid hormone in man. Proc. Soc. Exp. Biol. Med., *128*:501, 1968.
17. Reiss, E., Canterbury, J. M., and Egdahl, R. H.: Experience with a radioimmunoassay of parathyroid hormone in human sera. Trans. Assoc. Amer. Physicians, *81*:104, 1968.
18. Reiss, E., and Canterbury, J. M.: Primary hyperparathyroidism: Application of radioimmunoassay to differentiation of adenoma and hyperplasia and to preoperative localization of hyperfunctioning parathyroid glands. New Eng. J. Med., *280*:1381, 1969.
19. Sherwood, L. M., O'Riordan, J. L. H., Aurbach, G. D., and Potts, J. T., Jr.: Production of parathyroid hormone by nonparathyroid tumors. J. Clin. Endocrinol., 27:140, 1967.
20. Straus, F. H., II, and Paloyan, E.: The pathology of hyperparathyroidism. Surg. Clin. N. Amer., *49*:27, 1969.
21. Strott, C. A., and Nugent, C. A.: Laboratory tests in the diagnosis of hyperparathyroidism in hypercalcemic patients. Ann. Int. Med., 68:188, 1968.
22. Tashjian, A. J., Jr., Levine, L., and Munson, P. L.: Immunochemical identification of parathyroid hormone in nonparathyroid neoplasms associated with hypercalcemia. J. Exp. Med., *119*:467, 1964.
23. Watson, L.: Study of hypercalcemia in cancer. Quart. J. Med., *33*:525, 1964 (Abstract).
24. Widrow, S. H., and Levinsky, N.: Effect of parathyroid extract on renal tubular calcium reabsorption in the dog. J. Clin. Invest., *41*:2151, 1962.

2929 S. Ellis Avenue
Chicago, Illinois 60616

Some Basic Aspects of Bone Metabolism in Relation to Osteoporosis

JOHN L. SKOSEY, M.D., Ph.D.*

Osteoporosis is a disorder of bone characterized by a reduction in bone mass. All of the characteristic components of bone are lost proportionately, so that the percentage composition of osteoporotic bone approximates that of normal bone. Osteoporosis is differentiated from osteomalacia, in which the principal alteration is a decrease in the mineral content of bone.[46]

The clinical manifestations of osteoporosis include back pain and skeletal fractures, often resulting from minimal trauma. Roentgenograms of the spine demonstrate demineralization when the bone mineral content has been reduced by approximately 30 per cent.[40] Further loss of bone substance may result in "bone failure"[26] and the characteristic radiographic signs of vertebral compression fractures and fish vertebrae (biconcavity of the vertebrae resulting from compression of the softened bone by the intervertebral discs).

Bone has a relatively low cellularity in comparison with other tissues. The inorganic mineral and extracellular organic matrix of bone are responsible for its structural integrity. The inorganic mineral comprises approximately 65 per cent of the tissue by weight, and is primarily calcium phosphate in the form of hydroxyapatite and related crystalline structures. The organic matrix comprises the remainder of the tissue, and is primarily (90 to 95 per cent) collagen.[20]

Generalized osteoporosis occurs in a number of endocrine disorders, with certain dietary deficiencies, in some inherited diseases, and with malignant disease. Most cases are not associated with known etiological factors and are termed idiopathic, postmenopausal, or senile, depending upon the age of the patient involved.[40] Consideration of a number of the factors which control the metabolism of bone may help in understanding the development of osteoporosis.

* Arthritis Foundation Postdoctoral Fellow; Assistant Professor of Medicine, The University of Chicago, and Argonne Cancer Research Hospital (operated by the University of Chicago for the United States Atomic Energy Commission), Chicago, Illinois

GENERAL FACTORS

Age

The incidence of osteoporosis (by radiological and histological criteria) is greater in patients with fracture of the femoral neck than in a control population.[38] Thus, the incidence of hip fracture following minimal trauma (e.g., a fall beginning with the patient in a standing position) can be taken as a reflection of the incidence of osteoporosis,[6] although other factors (poor vision, incoordination, muscle weakness) cannot be disregarded.[3] Bollet et al.[6] estimated that the rate (per 1000 population per year) of hip fracture in white women due to minimal trauma increased from 2.0 at ages 50 to 64 to 5.0 at ages 65 to 74 to 16 at ages greater than 75 years. Bauer[3] similarly found that after age 50 the incidence of femoral neck fracture doubled in women with each 5-year increment in age. Using radiological criteria for osteoporosis, Gitman and Kamholtz[19] found that the incidence of osteoporosis in 933 women increased by 7.9 per cent for each 10-year increment in age over 65. Thus, the common clinical observation that osteoporosis is more common in the elderly is well documented.

Selle and Jurist[54] presented evidence suggesting that osteoporosis is not simply part of the uncomplicated aging process. They used an acoustical method to estimate the density of the ulna.[53] In brief, an audio generator producing a constant amplitude output was connected by a driver to the olecranon process of the ulna. A crystal transducer at the styloid process detected the input from the driver, as modified by the ulna. The frequency of the driver was altered to determine the frequency which produced the maximal output at the styloid process. This "resonant frequency" multiplied by the length of the ulna (FL) reflected the density of the ulna. The frequency distribution of FL in 59 clinically normal women over 45 years of age was not a normal distribution, but was bimodal. Thirty-seven per cent of this group of women had values of FL clustered in the lower range, and were suggested to be "pre-osteoporotic."

Activity

Osteoporosis occurs commonly following disuse or immobilization.[40] Such factors as diminished activity may contribute to the osteoporosis of aging, and weightlessness may contribute to the bone changes observed in subjects having undergone prolonged flights in space.[32] Although precise mechanisms have not been defined, physical factors such as tension and pressure modify the metabolism of connective tissue. Immobilization of canine forelimbs resulted in equal percentage losses of bone ash and collagen, as is the case in other forms of osteoporosis.[31] Increasing the venous pressure in canine tibia by means of a tourniquet resulted in a net increase in new and total bone.[34] Douglas et al.[11] studied collagen synthesis in healing flexor tendons of chickens by measuring the incorporation of radioactive proline into peptide-bound hydroxyproline. Hydroxyproline is present in vertebrates essentially only in collagen (except for a small amount in elastin), and is derived only from proline, which is hydroxylated after it is incorporated into the

peptides which form collagen.[43] The conversion of radioactive free proline to radioactive peptide-bound hydroxyproline serves as a measure of collagen synthesis. When chicken flexor tendons were incised and resutured, and then sectioned proximally to relieve tension on the original wound, wound healing and collagen synthesis (incorporation of ^{3}H-proline into ^{3}H-hydroxyproline) were impaired, as compared with that of tendon healing under tension.

Skin Changes in Osteoporosis

The skin of patients with osteoporosis is transparent.[41] The main feature of transparent skin is that the details of the veins are easily seen. Transparency correlates better with osteoporosis than reduction in skin fold thickness.

Collagen is the primary structural protein of skin as well as of bone. Collagen is a fibrous protein, each molecule consisting of three strands (three α-components) having a helical configuration and linked by *intra*molecular bonds. *Inter*molecular bonding links the collagen molecules together and contributes substantially to the strength of the connective tissues. Lathyrogens such as β-aminoproprionitrile (BAPN), interfere with the formation of intramolecular and intermolecular bonds.[7] Tissues of animals treated with BAPN have a markedly reduced tensile strength; that is, the force required to rupture a standard cross section of tissue is substantially less than that required to rupture normal tissue.[24] As part of the normal aging of skin, the degree of intermolecular cross linking increases, accompanied by an increase in the tensile strength of the skin.[45]

McConkey et al.[42] found that transparent skin had characteristic histological changes. The collagen bundles were finer and more widely dispersed in comparison with the coarse, compact collagen bundles of normal skin. The total amount of collagen was less than normal, and the skin was more hydrated. The rate of biosynthesis of protein (estimated by incorporation of ^{14}C-proline into skin in vitro) was at least equal to that of normal skin. They suggested that their observations could result from an impairment in the cross linking process. The resultant sparsely cross linked collagen would be more metabolically active and subject to loss from the skin. It is of interest that the reduced collagen content and increased hydration of transparent skin parallels the changes in these components in osteoporotic bone.[31]

Sex

The incidence of osteoporosis (using radiological criteria) in 933 women over age 65 was three times the incidence in a similarly aged group of 585 men.[19] The rates, for men and women, of hip fracture secondary to minimal trauma reflected this sex ratio.[6] Nordin et al.[47] studied 152 normal women (visitors to hospitalized patients) aged 19 to 83 years. They found that bone mass (estimated radiologically) decreased with age, and that the rate of decrease was accentuated after the menopause. Thus there is at least circumstantial evidence that loss of ovarian function is an important factor in the development of osteoporosis.

Jowsey et al.[30] used a microradiographic technique to measure the percentage of total surface of bone specimens involved in bone formation and bone resorption. The surface of bone from osteoporotic subjects involved in bone formation was normal, while that involved in bone resorption was increased. In adult mammals the hydroxyproline-containing peptides of urine derive mainly from mature, highly cross linked collagen. Bone is a potential source of urinary hydroxyproline. Katz and Kappas[33] found that estradiol and estriol each diminished urinary hydroxyproline excretion in man. Estrogen therefore may counteract the influence of factors which enhance bone resorption.

Race and Heredity

Osteoporosis appears to be significantly less common in Negroes than in whites. Bollet et al.[6] estimated the rate of hip fracture due to minimal trauma in white women over the age of 50 to be 2.9 per 1000 population per year in Charlottesville, Virginia, during the years 1960 to 1962. The corresponding rate for Negro women was 0.1. The rate for white men over age 50 was 0.8; no Negro men developed such hip fractures during the course of the study.

Studies of proline and hydroxyproline metabolism suggest deranged collagen metabolism in osteogenesis imperfecta. Proline and hydroxyproline (derived from proline incorporated into the peptide precursor of collagen) together account for approximately 22 per cent of the amino acid residues of bone collagen.[43] Studies of the metabolism of proline may therefore roughly reflect the synthesis of collagen, although proline, unlike hydroxyproline, is also found in other proteins. The blood half-life of injected proline was reduced in patients with osteogenesis imperfecta to approximately half that of pediatric control subjects.[59] Urinary peptide-bound proline[12, 59] and hydroxyproline[12, 37] were also increased. These studies suggested that both the synthesis and degradation of collagen are enhanced in this disorder, the net result being a diminished amount of collagen.

Liakakos et al.[38] presented evidence suggesting that the osteoporosis observed in thalassemia may be due to a metabolic defect in collagen synthesis and not simply secondary to increased intramedullary pressure resulting from hyperplastic bone marrow. They found proline levels to be elevated in both plasma and urine, and suggested that this reflected a defect in the utilization of the amino acid for collagen synthesis. Finerman et al.[14] found that aminoisobutyric acid, L-azetidine-2-carboxylic acid, and hydroxyproline each interfered with the entry of proline into bone cells in vitro, and that collagen synthesis was subsequently inhibited. Factors which diminish transport of constituents necessary for collagen synthesis may contribute to decreased production of the protein and reduced bone mass.

NUTRITIONAL FACTORS

General

Osteoporosis is observed in those suffering from generalized malnutrition.[40] Fasting serves as a model of nutritional deprivation. Obese

subjects, when fasted, demonstrated changes in urinary excretion of both calcium and hydroxyproline.[39] Calcium excretions rose during the period of metabolic acidosis which accompanied fasting and returned toward normal before the end of the fast. Urinary hydroxyproline rose early in the fast and remained high even into the period of re-feeding. This study suggested that collagen of bone and other tissues may be utilized to fulfill nutritional requirements during fasting or, alternatively, that nutritional factors are required to maintain collagen and inhibit its resorption.

As noted above, the structural integrity of collagen depends largely upon its intermolecular cross links. Deyl et al.[10] found reduced cross linking in the skin of rats fed 60 per cent as much food as control animals. Similar changes in bone would be expected to alter its structure adversely.

Ascorbic Acid

The relationship of ascorbic acid deficiency to osteoporosis is illustrated by the study of Seftel et al.[52] They reported 32 cases of "osteoporosis, scurvy, and siderosis in Johannesburg Bantu," a syndrome of special interest in that the patients were Negro, predominantly male (26 of the 32 cases) and relatively young (the majority between 40 and 55 years old). An enzyme isolated from chick embryo extracts catalyzes the hydroxylation of peptide-bound proline of collagen precursors. This "protocollagen hydroxylase" requires ascorbic acid, ferrous ion, and 2-ketoglutarate as cofactors.[50] Jeffrey and Martin[29] have found that bone collagen synthesis similarly requires ascorbic acid. Embryonic chick tibia grown in tissue culture, when deprived of ascorbic acid, fails to make collagen.[28] The synthesis of other proteins, deoxyribonucleic acid, and ribonucleic acid is also impaired, although less drastically than collagen synthesis. The tibia do not grow at a normal rate. Normal collagen synthesis may be a prerequisite for normal bone growth; alternatively, biochemical processes other than collagen synthesis may require ascorbic acid.

Calcium

Bone is constantly undergoing resorption and regeneration. It is obvious that bone cannot be normally produced, nor its structure maintained, without an adequate supply of mineral which is crystallized normally upon the organic matrix. Without entering the controversy as to whether osteoporosis results from a defect primarily in the mineral or organic phase of bone, it should be safe to suggest that osteoporosis does not result simply from lack of dietary calcium. While recognizing the limitations of retrospective dietary histories, two groups of workers could find no differences in the intakes of calcium by osteoporotic and nonosteoporotic individuals.[4, 55] Avioli et al.[2] found that patients with osteoporosis absorbed calcium normally when compared with age-matched control subjects. Although it is easier to devise and interpret experiments which consider mineral metabolism separate from matrix (collagen) metabolism, the following illustrations suggest that this approach may be oversimplified.

As noted above, the degree of cross linking in collagen determines a number of the physical characteristics of the protein. Newly formed collagen contains few cross links. As the collagen "matures," it becomes more highly cross linked and insoluble. A small percentage of the total collagen of most tissues has a small enough number of cross links to be "soluble" – that is, extractable in 2.0 M sodium chloride, or in dilute acid. (Bone collagen represents an exception to this generalization. Only about 1 per cent of bone collagen can be extracted by these mild procedures.[43] Calcium-collagen interactions, described below, may contribute to this relative decrease in soluble collagen in bone.) The solubility of a given collagen fraction is then inversely proportional to its age. As noted earlier, "age" of collagen is also used in a slightly different context. In most tissues, the percentage of mature, insoluble collagen increases with the age of the animal.[45]

Engström[13] studied the x-ray diffraction pattern of noncalcified and calcified turkey tendon. Wet tendon (calcified or noncalcified) exhibited the 660 Å spacing characteristic of collagen. Upon drying, the period of the noncalcified tendon shortened, whereas that of the calcified tendon remained constant. He concluded that a moderate apatite deposition stabilized the collagen, making it resistant to shrinkage during dehydration.

The results of other studies similarly indicate that calcium contributes to the stability of the collagen structure. Hruza and Hlavackova[27] studied the contraction-relaxation curves in sodium perchlorate of collagen from tail tendon of young and old rats. When the old rat tendon collagen was decalcified (by treatment with EDTA, oxalate, or citrate) it behaved as collagen from young rats. This change was reversed by the addition of $CaCl_2$. Young rat tendon collagen behaved as that from older animals upon the addition of $CaCl_2$. Steven[57] found that treatment of bovine and human tendons with chelating agents (EDTA, ethylenediamine, salicylate, D-penicillamine) greatly increased the amount of tendon collagen which could be dispersed in dilute acetic acid. This dispersibility was reversed by treatment with calcium acetate.

Naber et al.[41] produced lathyrism in chick embryos by incubating the embryos with BAPN. They observed the characteristic increase in soluble collagen resulting from impaired cross linking. Incubation with calcium gluconate counteracted this effect of BAPN; the amount of collagen soluble in 1.0 M sodium chloride from embryos incubated with both BAPN (0.4 mg. per egg) and calcium gluconate (15 mg. per egg) was decreased almost to that of control embryos. In addition, they observed that BAPN inhibited the incorporation of proline-^{14}C into hydroxyproline-^{14}C of insoluble collagen (the fraction remaining after serial extractions with 0.14 M NaCl, 1.0 M NaCl, and citrate buffered at pH 3.5). The addition of calcium gluconate resulted in incorporation of ^{14}C approaching the control value.

Fluoride

Fluoride given in treatment of osteoporosis is incorporated into bone mineral and the resultant crystals are larger than those from normal or

osteoporotic bone.[23] Clinical improvement (such as reduction of back pain) and increase in bone density are observed.[8, 9] It is disturbing, therefore, that the administration of fluoride to rodents either does not improve bone strength[51] or results in decreased breaking strength.[18] Comparison of populations living in areas having high and low fluoride contents of drinking water suggest that these studies cannot be extrapolated to man. Ansell and Lawrence[1] studied two populations in Great Britain, one living in an area having a water supply with 1.0 part per million of fluoride, and the other with 0.2 part per million of fluoride. The rate of osteoporosis assessed by estimation of bone density was lower in the "high fluoride" area. Unfortunately, their study did not include an estimation of the frequency of collapsed vertebrae or other measure of "bone failure". Bernstein et al.[4] studied two populations in North Dakota with high (4 to 5.8 parts per million) and low (0.15 to 0.3 part per million) drinking water fluoride contents. Although the incidence of collapsed vertebrae in men was high in both groups and not significantly different, the women in the high fluoride population had significantly fewer collapsed vertebrae. Thus, on balance, the evidence suggests that fluoride may be useful in preventing and reversing changes of osteoporosis, although this question cannot be considered settled.

ENDOCRINE FACTORS

Glucocorticoids

Björk[5] estimated the mineral content of the ulna by making roentgenograms of the bone and an aluminum wedge (ranging in thickness from 0.5 to 7 mm.) simultaneously. He pointed out that because of the wide variation among individuals this test could not be used to diagnose osteoporosis. It was, however, useful in following bone density changes in a given person. He found that two individuals with osteoporosis secondary to steroid treatment had marked increases in bone mineral content 2½ and 3½ months after the steroid dosage was reduced. Jowsey et al.[30] inferred from their estimates of percentage of bone surface involved in formation and resorption that resorption was increased and formation diminished in the osteoporosis resulting from glucocorticoid excess. The return of both of these processes toward normal following cessation of steroid therapy may account for the restoration of bone observed by Björk.[5] Although it is generally agreed that methods of preventing osteoporosis are superior to those used for treatment,[40] there is hope for some reversal in the osteoporotic process following removal of this offending agent.

The results of animal experiments are consistent with the concept of a bimodal action of glucocorticoids upon bone and other connective tissue. Smith and Allison[56] measured this incorporation of glycine-^{14}C (which accounts for about one third of the amino acid residues of collagen) into the femur total collagen of weanling, young adult, and adult rats. Pretreatment of the animals with cortisone acetate for 7 days prior to injection of glycine-^{14}C resulted in a 50 per cent decrease in incorpora-

tion of the label into collagen. Peck, Brandt, and Miller[49] found that hydrocortisone inhibited the incorporation of proline-^{14}C into hydroxyproline-^{14}C of isolated bone cells maintained in culture. The incorporation of proline-^{14}C into noncollagenous protein and of uridine-^{14}C into ribonucleic acid were also inhibited, indicating that the effects of hydrocortisone were not limited to collagen metabolism.

Houck and Patel[25] found that 2 hours after the administration of either fluoroprednisolone or hydrocortisone to rats, the salt-soluble collagen of skin decreased, the citrate-extractable collagen increased, and insoluble collagen decreased. When the animals were given proline-^{3}H along with the hydrocortisone, the specific activity of the citrate extractable collagen of the treated animals was only one fifth that of the controls. Their interpretation was that steroid treatment rendered the insoluble collagen more acid-extractable. They further found that glucocorticoid treatment resulted in the appearance of a protease (measured by its digestion of denatured hemoglobin). In addition, they demonstrated a collagenase zymogen which could be activated (to attack collagen) by limited digestion with trypsin. They proposed that the glucocorticoids cause release of a cellular protease which partially digests (and thereby activates) a collagenase zymogen. The active collagenase then digests collagen. Although their studies were done on skin, their proposed mechanism could explain the increased bone resorption observed with glucocorticoid treatment.

Insulin

Selle and Jurist,[54] using the acoustical technique described above to estimate bone mass, found decreases in patients with diabetes. Although factors other than insulin deficiency (whether absolute or relative) are certainly operative in diabetes mellitus, the study of Wettenhall et al.[60] may bear on this observation. They maintained tibias from newborn rats in a continuous-flow bone culture system. Insulin added to the culture medium stimulated the incorporation of proline-^{3}H into bone hydroxyproline-^{3}H (collagen synthesis) and into bone peptide-bound proline-^{3}H (total protein synthesis).

Thyroid Hormone

The urinary excretion of peptide-bound hydroxyproline increased in rats made hyperthyroid by thyroxine administration, and decreased when hypothyroidism was induced by methimazole.[15] Kivirikko et al.[35] found that the increased hydroxyprolinuria in hyperthyroid rats resulted from increased rates of destruction of both soluble (newly formed) and insoluble (older) collagen. When they administered proline-^{14}C to hyperthyroid rats, the amount and specific activity of urinary hydroxyproline were increased early during the time when the label was present as hydroxyproline-^{14}C only in the soluble collagen of tissue. When rats were given proline-^{14}C a month before the beginning of thyroxine treatment, the label was found only in tissue-insoluble collagen. Administration of thyroxine again resulted in increased amount and specific activity of urinary hydroxyproline. Kivirikko et al.[35] also found that, in hyper-

thyroidism, incorporation of proline-^{14}C to skin collagen hydroxyproline-^{14}C was diminished. Thus, as in the case with glucocorticoids, thyroxine may decrease bone mass by both enhancing destruction and decreasing formation of the tissue.

Parathyroid Hormone

As pointed out by Flanagan and Nichols,[16] hyperparathyroidism has a number of clinical manifestations, including generalized osteoporosis. They studied the metabolism, in vitro, of bone obtained at biopsy from five patients with hyperparathyroidism due to "continuous" hypersecretion of parathormone (characterized by hypercalcemia, elevated serum alkaline phosphatase, hypercalciuria, hyperhydroxyprolinuria, and evidence of bone disease). Bone from these patients, compared with that from normal control subjects, had an increased oxygen uptake, lactate production, collagen synthesis (incorporation of proline-^{14}C into the collagen fraction), and glucose incorporation into "cell fractions" (noncollagenous proteins, alkali-soluble collagen and mucopolysaccharide, and deoxyribonucleic acid).

Laitinen,[36] studying the effect of parathyroid hormone upon the metabolism of proline-^{14}C and $^{45}CaCl_2$ by rats, similarly found that parathormone stimulated both synthesis and degradation of bone. The animals were treated for 8 days with parathyroid extract, and then injected with proline-^{14}C and $^{45}CaCl_2$. Compared with control animals, parathyroid treated animals exhibited increased incorporation of proline-^{14}C into bone collagen hydroxyproline-^{14}C, and increased accretion of ^{45}Ca into bone. The specific activity and total amount of urinary hydroxyproline increased, indicating enhanced breakdown of newly formed collagen by the parathyroid-treated animals. When animals were given proline-^{14}C and $^{45}CaCl_2$ a month before institution of treatment with parathyroid extract (the labels were thus fixed in mature collagen and hydroxyapatite), the hormone again enhanced specific activity of urinary hydroxyproline-^{14}C and ^{45}Ca. Thus, parathyroid extract enhanced bone formation and degradation of newly formed and older bone.

Heparin potentiates the stimulation of bone resorption by parathyroid hormone. This observation may bear on some forms of osteoporosis. Griffith et al.[22] reported six patients who had received 15,000 to 30,000 units of heparin daily for 6 months or longer, and who developed fractures of the ribs or vertebrae either spontaneously or after minimal trauma. Goldhaber[21] studied calvarias from newborn mice maintained in tissue culture. Neither heparin (10 units per ml.) nor parathyroid extract (0.1 unit per ml.) caused bone resorption in the concentrations used. The two agents in combination, however, caused extensive bone resorption to an extent usually seen with parathyroid extract at a concentration of 0.5 units per ml. Frame and Nixon[17] found an increased number of mast cells (a natural source of heparin) in the bone marrow of patients with osteoporosis. These studies illustrate that factors which alone have no detectable influence on bone metabolism may, in combination, have profound effects. They point out one problem in the search for causes of osteoporosis.

SUMMARY

Bone is constantly being resorbed and regenerated. Osteoporosis can result from increased resorption, decreased regeneration, or a combination of these processes. Alterations in rates of resorption and regeneration can result from alterations in metabolism at a number of sites. Consideration of factors which control bone metabolism help to explain the osteoporosis associated with some diseases. The bulk of cases of osteoporosis are as yet inadequately explained and must be termed "idiopathic" or by some other nondescript term. Further knowledge of the factors which control bone metabolism will help to remove increasing numbers of cases of osteoporosis from the "idiopathic" group to a better defined category.

ACKNOWLEDGMENT

I thank Mrs. Genevieve LaPinska for assistance in preparation of this paper.

REFERENCES

1. Ansell, B. M., and Lawrence, J. S.: Fluoridation and the rheumatic diseases. A comparison of rheumatism in Watford and Leigh. Ann. Rheum. Dis., *25*:67, 1965.
2. Avioli, L. V., McDonald, J. E., and Lee, S. W.: The influence of age on the intestinal absorption of ^{47}Ca in women and its relation to ^{47}Ca absorption in postmenopausal osteoporosis. J. Clin. Invest., *44*:1960, 1965.
3. Bauer, G. C. H.: Epidemiology of fracture in aged persons. Clin. Orthop., *17*:219, 1960.
4. Bernstein, D. S., Sadowsky, N., Hegsted, D. M., Guri, C. D., and Stare, F. J.: Prevalence of osteoporosis in high- and low-fluoride areas in North Dakota. J.A.M.A., *198*:499, 1966.
5. Björk, L.: Radiographic determination of the bone mineral content in osteoporosis. Acta Radiol., *3*:218, 1965.
6. Bollet, A. J., Engh, G., and Parson, W.: Epidemiology of osteoporosis. Arch. Int. Med., *116*:191, 1965.
7. Bornstein, P., Kang, A. H., and Piez, K. A.: The nature and location of intramolecular cross-links in collagen. Proc. Nat. Acad. Sci., *55*:417, 1966.
8. Cass, R. M., Croft, J. D., Jr., Perkins, P., Nye, W., Waterhouse, C., and Terry, R.: New bone formation in osteoporosis following treatment with sodium fluoride. Arch. Int. Med., *118*:111, 1966.
9. Cohen, P., and Gardner, F. H.: Induction of skeletal fluorosis in two common demineralizing disorders. J.A.M.A., *195*:962, 1966.
10. Deyl, Z., Everitt, A., and Rosmus, J.: Changes in collagen cross-linking in rats with reduced food intake. Nature, *217*:670, 1968.
11. Douglas, L. G., Jackson, S. H., and Lindsay, W. K.: The effects of dexamethasone, norethandrolone, promethazine and a tension-relieving procedure on collagen synthesis in healing flexor tendons as estimated by tritiated proline uptake studies. Canad. J. Surg., *10*:36, 1967.
12. Dubovsky, J., Pacovsky, V., and Silinkova-Malkova, E.: Die Ausscheidung der Hydroxyprolin- und Prolinpeptide bei Osteomalazie und Osteoporose. Zschr. Inn. Med., *21*:231, 1966.
13. Engström, A.: Apatite-collagen organization in calcified tendon. Exp. Cell Res., *43*:241, 1966.
14. Finerman, G. A. M., Downing, S., and Rosenberg, L. E.: Amino acid transport in bone. II. Regulation of collagen synthesis by perturbation of proline transport. Biochim. Biophys. Acta, *135*:1008, 1967.
15. Fink, C. W., Ferguson, J. L., and Smiley, J. D.: Effect of hyperthyroidism and hypothyroidism on collagen metabolism. J. Lab. Clin. Med., *69*:950, 1967.
16. Flanagan, B., and Nichols, G., Jr.: Metabolic studies of human bone *in vitro*. II. Changes in hyperparathyroidism. J. Clin. Invest., *44*:1795, 1965.

17. Frame, B., and Nixon, R. K.: Bone-marrow mast cells in osteoporosis of aging. New Eng. J. Med., *279*:626, 1968.
18. Gedalia, I., Frumkin, A., and Zukerman, H.: Effects of estrogen on bone composition in rats at low and high fluoride intake. Endocrinology, *75*:201, 1964.
19. Gitman, L., and Kamholtz, T.: Incidence of radiographic osteoporosis in a large series of aged individuals. J. Gerontol., *20*:32, 1965.
20. Glimcher, M. J., and Krane, S. M.: The organization and structure of bone, and the mechanism of calcification. *In* Gould, B. S., ed.: Treatise on Collagen. New York, Academic Press, 1968, Vol. 2, part B, p. 67.
21. Goldhaber, P.: Heparin enhancement of factors stimulating bone resorption in tissue culture. Science, *147*:407, 1965.
22. Griffith, G. C., Nichols, G., Jr., Asher, J. D., and Flanagan, B.: Heparin osteoporosis. J.A.M.A., *193*:91, 1965.
23. Gron, P., McCann, H. G., and Bernstein, D.: Effect of fluoride on human osteoporotic bone mineral. J. Bone Joint Surg., *48-A*:892, 1956.
24. Harkness, R. D.: Mechanical properties of collagenous tissues. *In* Gould, B. S., ed.: Treatise on Collagen. New York, Academic Press, 1968, Vol. 2, part A, p. 247.
25. Houck, J. C., and Patel, Y. M.: Proposed mode of action of corticosteroids on the connective tissue. Nature, *206*:158, 1965.
26. Howell, D. S.: Metabolic bone diseases. *In* Hollander, J. L., ed.: Arthritis and Allied Conditions. Philadelphia, Lea & Febiger, 1966, p. 973.
27. Hruza, Z., and Hlavackova, V.: Effect of decalcification on chemical contraction and relaxation of collagen fibres. Gerontologia, *13*:246, 1967.
28. Jeffrey, J. J., and Martin, G. R.: The role of ascorbic acid in the biosynthesis of collagen. I. Ascorbic acid requirement by embryonic chick tibia in tissue culture. Biochim. Biophys. Acta, *121*:269, 1966.
29. Jeffrey, J. J., and Martin, G. R.: The role of ascorbic acid in the biosynthesis of collagen. II. Site and nature of ascorbic acid participation. Biochim. Biophys. Acta, *121*:281, 1966.
30. Jowsey, J., Kelly, P. J., Riggs, B. L., Bianco, A. J., Jr., Scholz, D. A., and Gershon-Cohen, J.: Quantitative microradiographic studies of normal and osteoporotic bone. J. Bone Joint Surg., *47-A*:785, 1965.
31. Kanefield, D. G., Klein, L., and Heiple, K. G.: Bone matrix in disuse osteoporosis. Surg. Forum, *17*:454, 1966.
32. Karpovich, P. V.: Exercise in medicine: A review. Arch. Phys. Med., *49*:66, 1968.
33. Katz, F. H., and Kappas, A.: Influence of estradiol and estriol on urinary excretion of hydroxyproline in man. J. Lab. Clin. Med., *71*:65, 1968.
34. Kelly, P. J.: Effect of unilateral increased venous pressure on bone remodeling in canine tibia. J. Lab. Clin. Med., *72*:410, 1968.
35. Kivirikko, K. I., Laitinen, O., Aer, J., and Halme, J.: Metabolism of collagen in experimental hyperthyroidism and hypothyroidism in the rat. Endocrinology, *80*:1051, 1967.
36. Laitinen, O.: Parathyroid-induced changes in collagen and calcium metabolism *in vivo*. Endocrinology, *80*:815, 1967.
37. Langness, U., and Behnke, H.: Die Hydroxyprolinausscheidung im Urin bei der Osteogenesis imperfecta. Klin. Wschr., *44*:1294, 1966.
38. Liakakos, D., Karponzas, J., and Agathopoulos, A.: Hyperprolinemia and hyperprolinuria in thalassemia. J. Pediat., *73*:419, 1968.
39. Litvak, J., Wortsman, J., Pumarino, H., and Armendaris, R.: Effect of fasting on urinary calcium and hydroxyproline excretion in obesity. J. Clin. Endocrinol., *28*:311, 1968.
40. Lutwak, L.: Nutritional aspects of osteoporosis. J. Amer. Geriat. Soc., *17*:115, 1969.
41. McConkey, B., Bligh, A. S., Fraser, G. M., and Whiteley, H.: Transparent skin and osteoporosis. Lancet, *1*:693, 1963.
42. McConkey, B., Walton, K. W., Carney, S. A., Lawrence, J. C., and Ricketts, C. R.: Significance of the occurrence of transparent skin. A study of histological characteristics and biosynthesis of dermal collagen. Ann. Rheum. Dis., *26*:219, 1967.
43. Miller, E. J., and Martin, G. R.: The collagen of bone. Clin. Orthop., *59*:195, 1968.
44. Naber, E. C., Scott, K., and Johnson, R. M.: Effect of calcium on the incorporation of proline into hydroxyproline of collagen fractions from normal and lathyritic chick embryos. Poult. Sci., *44*:1540, 1965.
45. Nimini, M. E., Guia, E. de, and Bavetta, L. A.: Collagen, hexosamine and tensile strength of rabbit skin during aging. J. Invest. Derm., *47*:156, 1966.
46. Nordin, B. E. C.: Osteoporosis and calcium deficiency. *In* Rohdahl, K., Nicholson, J. T., and Brown, E. M., Jr., eds.: Bone as a Tissue. New York, McGraw-Hill Book Co., 1960, p. 46.
47. Nordin, B. E. C., MacGregor, J., and Smith, D. A.: The incidence of osteoporosis in normal women: Its relation to age and the menopause. Quart. J. Med., N.S., *35*:25, 1966.
48. Peck, W. A., Birge, S. J., Jr., and Brandt, J.: Collagen synthesis by isolated bone cells: Stimulation by ascorbic acid *in vitro*. Biochim. Biophys. Acta, *142*:512, 1967.

49. Peck, W. A., Brandt, J., and Miller, I.: Hydrocortisone-induced inhibition of protein synthesis and uridine incorporation in isolated bone cells in vitro. Proc. Nat. Acad. Sci., 57:1599, 1967.
50. Prockop, D., and Kivirikko, K. I.: Hydroxyproline and the metabolism of collagen. *In* Gould, B. S., ed.: Treatise on Collagen. New York, Academic Press, 1968, Vol. 2, part A, p. 215.
51. Saville, P. D.: Water fluoridation: Effect on bone fragility and skeletal calcium content in the rat. J. Nutrition, *91*:353, 1967.
52. Seftel, H. C., Malkin, C., Schmaman, A., Abrahams, C., Lynch, S. R., Charlton, R. W., and Bothwell, T. H.: Osteoporosis, scurvy, and siderosis in Johannesburg Bantu. Brit. Med. J., *1*:642, 1966.
53. Selle, W. A., and Jurist, J. M.: Acoustical detection of senile osteoporosis. Proc. Soc. Exp. Biol. Med., *121*:150, 1966.
54. Selle, W. A., and Jurist, J. M.: The onset of postmenopausal osteoporosis as studied by a new technique. J. Amer. Geriat. Soc., *14*:930, 1966.
55. Smith, R. W., and Frame, B.: Concurrent axial and appendicular osteoporosis. New Eng. J. Med., *273*:73, 1965.
56. Smith, Q. T., and Allison, D. J.: Skin and femur collagens and urinary hydroxyproline of cortisone-treated rats. Endocrinology, *77*:785, 1965.
57. Steven, F. S.: The effect of chelating agents on collagen interfibrillar matrix interactions in connective tissue. Biochim. Biophys. Acta, *140*:522, 1967.
58. Stevens, J., Freeman, P. A., Nordin, B. E. C., and Barnett, E.: The incidence of osteoporosis in patients with femoral neck fracture. J. Bone Joint Surg., *44-B*:520, 1962.
59. Summer, G. K., and Patton, W. C.: Intravenous proline tolerance in osteogenesis imperfecta. Metabolism, *17*:46, 1968.
60. Wettenhall, R. E. H., Schwartz, P. L., and Bornstein, J.: Actions of insulin and growth hormone on collagen and chondroitin sulfate synthesis in bone organ cultures. Diabetes, *18*:280, 1969.

950 East 59th Street
Chicago, Illinois 60637

Practical Diagnostic Aids in the Study of Dyslipoproteinemias

ANGELO M. SCANU, M.D.*

The purpose of this article is to provide a brief survey of the methods currently utilized in the clinical laboratory for the study of dyslipoproteinemias, disorders characterized by an abnormal distribution of plasma lipoproteins. The physicochemical basis for the classification of such disorders and for the methods for their detection has been presented in detail in a number of review articles. The most pertinent reviews are listed for the convenience of the reader in Table 1.

Table 1. *Significant Review Articles on Plasma Lipoproteins*

TOPIC	AUTHORS	YEAR
Physicochemistry	Scanu[15]	1965
Physicochemistry	Nichols[10]	1967
Physicochemistry	Schumaker and Adams[18]	1969
Metabolism and clinical	Fredrickson et al.[4]	1967
Methodology	Hatch and Lees[6]	1968

Comparison Between Ultracentrifugation and Electrophoresis

Although preparative and analytical ultracentrifugation remain the best tools for the separation and quantification of plasma lipoproteins, based on their hydrated density, appropriate electrophoretic techniques afford a fractionation of these complexes which is satisfactory for clinical purposes. This is corroborated by the comparative parameters reported in Table 2.

Comments on Electrophoretic Techniques

The technique of paper electrophoresis by Lees and Hatch[7] employing albuminated buffers has been extensively utilized for the qualitative separation of serum lipoproteins and has represented the basis for the

* Associate Professor of Medicine and Biochemistry, The University of Chicago, and Argonne Cancer Research Hospital (operated by the University of Chicago for the United States Atomic Energy Commission), Chicago, Illinois. Recipient of USPHS Research Career Development Award HE-24867

Table 2. *Classification of Plasma Lipoproteins Based on Their Hydrated Density (Ultracentrifugation) and Electric Charge Density (Electrophoresis)*

ULTRACENTRIFUGATION		ELECTROPHORESIS*
HYDRATED DENSITY (gm. per ml.)	NOMENCLATURE	
< 0.93	Chylomicrons	–
0.93–1.01	VLDL	pre-β
1.01–1.05	LDL	β
1.08–1.16	HDL**	α_1

VLDL, Very low density lipoprotein; LDL, low density lipoprotein; HDL, high density lipoprotein.

*This refers to anodic migration at alkaline pH in various supporting media, such as paper, cellulose acetate, agar, or agarose.

**For simplicity this lipoprotein will be considered as a single class.

classification of dyslipoproteinemias by Fredrickson and Lees.[3] However, other electrophoretic techniques have been described which appear to possess a sufficient degree of simplicity and reproducibility to warrant extensive testing. Of particular significance are those employing as a supporting medium cellulose acetate,[2] starch gel,[8] acrylamide gel,[9] and agarose gel.[11] Of these, the latter technique, described by Noble,[11] is particularly recommendable because of its relative simplicity, good resolving power, and the possibility of application on a semiquantitative level. Recently, it has been shown that the results of lipoprotein analysis by this method compare well with those obtained by analytical ultracentrifugation.[12]

Basic Composition of Plasma Lipoproteins

In order to correlate the results of the electrophoretic technique with those obtained by the chemical analyses of serum lipids, one must be acquainted with the chemical composition of each lipoprotein constituent. These data are given in Table 3. One can easily observe that there are remarkable differences in the chemical composition of each lipoprotein class, both in terms of protein:lipid ratio and lipid distribution. These lipoproteins differ also in the nature of their protein moiety and this can account for the differences in their electrical charge density and immunological specificity.

Combined Application of Electrophoretic and Chemical Techniques to the Diagnosis of Hyperlipoproteinemias

In this discussion, it is assumed that the reader is already familiar with the typing system of hyperlipoproteinemia (types I to V) proposed by Fredrickson and Lees.[3] The essentials for such diagnosis based on electrophoretic and chemical analyses are given in Table 4. Two things clearly emerge from this table: one, the close correlation between electrophoretic and chemical results as predicted by the data in Table 3; two, the parallel behavior of serum total cholesterol and phospholipids.

Table 3. *Chemical Composition of the Major Lipoproteins of Human Plasma*

	CHYLOMICRONS	PRE-β	β	α
Protein	2	10	23	50
Cholesterol-free	2	5	10	2
Cholesterol-ester	4	13	36	20
Phospholipid	5	20	21	24
Triglycerides	87	52	10	4

It is apparent, therefore, that in the diagnosis of hyperlipoproteinemias we may dispense with one of the two chemical determinations and confine our analysis to serum cholesterol (which is relatively easy to determine) and use the results in conjunction with serum triglycerides. Two questions may be asked at this point:

1. *Is electrophoresis alone sufficient for the diagnosis of hyperlipoproteinemias?* In general, the answer to this question is no, and this is primarily because electrophoresis on either paper or other supportive medium is, at best, semiquantitative and does not provide an accurate indication of the degree of increase or decrease of any of the plasma lipoproteins. It may also be added that in type III hyperlipoproteinemia, attributed to the presence in plasma of a β-lipoprotein with unusual affinity for triglycerides, there is need for complementary studies with ultracentrifuge to separate the abnormal lipoproteins by differential flotation. On the other hand, electrophoresis alone may suggest that appropriate diagnosis in cases where there is clear evidence for either a chylomicron or a pre-β band, or in cases where one of the major lipoprotein bands is not detected. Even in these instances, however, chemical analyses are still required.

2. *Is lipid analysis alone sufficient for diagnosis of hyperlipoproteinemias?* The answer may be yes for hyperlipoproteinemias types

Table 4. *Essential Laboratory Features of Hyperlipoproteinemias*

HYPER-LIPOPROTEINEMIA TYPE*	PREDOMINANT BAND BY ELECTROPHORESIS	PRINCIPAL ELEVATION OF SERUM LIPIDS		
		TRIGLYCERIDES	CHOLESTEROL	PHOSPHOLIPIDS
I	Chyl	+++	–	–
II	β		+++	+++
III	β (broad)	++	++	++
IV	Pre-β	++	+	+
V	Chyl-pre β-β	+++	+++	+++

*According to Fredrickson et al.[4]

I and II, but it becomes equivocal for the other types. In the case of type I, where there is a marked elevation of serum chylomicrons (see Table 4) and consequently of triglycerides (see Table 3) the serum, according to the degree of hyperlipemia, appears opalescent, turbid, milky, or creamy. A practical hint in these cases is to leave such a serum to sit in the cold overnight. Because of the very high flotational rates of these particles, they can be expected to rise to the top of the tube and leave a perfectly clear subnatant. A chemical analysis of this top fraction should now be compatible with that reported for chylomicrons in Table 3. This and the added documentation of a deficiency of lipoprotein lipase should now be sufficient for the diagnosis of type I hyperlipoproteinemia. In turn, type II hyperlipoproteinemia is to be suspected when dealing with clear sera with a high concentration of cholesterol and essentially normal levels of triglycerides. One anticipates here an intensive β band by electrophoresis, a normal α, and absence of pre-β.

From the above, we may conclude that in general both electrophoretic and chemical data are needed for reaching an accurate classification of hyperlipoproteinemia, and that these techniques may have to be complemented occasionally by either ultracentrifugation or enzymatic studies. In addition, one needs an accurate clinical evaluation of the case, which may favor differentiation between primary or secondary forms by establishing possible disease states associated with or the consequence of hyperlipoproteinemia, and also the type of familial incidence of the disease.

Application of Electrophoretic and Chemical Techniques to the Diagnosis of Hypolipoproteinemia (Table 5)

In contrast to hyperlipoproteinemias, hypoliproteinemias as might be expected are characterized by low levels of plasma lipids. Only in homozygous patients does one expect to observe total absence of either α or β lipoproteins. It would appear, however, that this is never the case, since in both a-β and a-α lipoproteinemia, abnormal β or α lipoproteins in significant levels have been detected immunologically—probably an expression of defective synthesis. As in hyperlipoproteinemia, these cases may be primary or secondary in origin and require for their diagnosis both clinical and laboratory tools. For a more detailed discussion the reader is referred to the article by Fredrickson et al.[1] It is of interest

Table 5. *Expected Laboratory Features of Hypolipoproteinemia*

HYPOLIPOPROTEINEMIA	ABNORMALITIES	
	ELECTROPHORETIC PATTERN	PLASMA LIPIDS
a-β lipoproteinemia	absent β	low
Hypo β-lipoproteinemia	trace β	low
a-α lipoproteinemia	absent α	low
Hypo α-lipoproteinemia	trace α	low

to mention the existence of a peculiar syndrome recently observed in three patients with familial plasma lecithin-cholesterol acyltransferase deficiency (the enzyme controls the reaction: lecithin + free cholesterol = lysolecithin + ester cholesterol) and characterized by decreased levels of plasma cholesterol esters, decreased amounts of lysolecithin, absence of α-lipoprotein band by paper or cellulose acetate electrophoresis, proteinuria, anemia and marked corneal opacity.[13] Progressive renal failure is a typical course of the disease. In these cases, separation of serum cholesterol into its free and esterified fractions is essential for the diagnosis and this may easily be achieved by thin layer chromatography on lipid extracts from whole serum. The three patients reported in the literature were found to have increased amounts of very low density serum lipoproteins, which has been considered as a secondary manifestation of the nephrotic syndrome present in these cases.

The Problem of Severe Hyperlipemias – Role of Immunology as a Diagnostic Aid

In heavy hyperlipemias, lipoproteins may be difficult to separate according to the normal criteria of charge density (electrophoresis) or hydrated density (ultracentrifugation). In such instances, it may become desirable to identify the various serum lipoprotein classes according to the nature of their protein moieties, which as stated earlier exhibit immunological specificity. Antisera to both serum β and α-lipoproteins are now commercially available (purity of the products should, however, be tested) or can easily be prepared in the laboratory and stored for prolonged periods. While the reader is referred for the technical details to a recent discussion on the subject,[6] mention should be made of the advantages of carrying out immunological studies on sera from which a large complement of lipids has been removed by extraction with organic solvents under appropriate experimental conditions.[4, 15] Such techniques have been valuable both in aiding structural studies of serum lipoproteins and also in defining their distribution (at least in terms of their protein moiety) where electrophoretic and ultracentrifugal techniques alone or in combination were not sufficient to provide unequivocal answers. Immunological techniques are expected to play an even greater role now that a more extensive knowledge of the physicochemical properties of the lipoproteins has been obtained in various laboratories and new subfractions are being recognized. A still unresolved problem is that of the suspected haptenicity of the lipid moiety, but answers to this problem are likely to be soon forthcoming.

The Significance of Typing Plasma Lipid Abnormalities as Dyslipoproteinemias

Since essentially all plasma lipids do not circulate free but in bound form as part of specific lipoprotein complexes, it is evident why lipid disorders should be more properly classified as lipoprotein disorders. This concept is corroborated by the observation that elevation or decrease of plasma lipids is associated with a parallel elevation or decrease of the protein moieties to which they are specifically bound. We may add

that there is increasing evidence that the protein moieties play a major structural role in these lipoproteins, suggesting that at least in certain conditions lipid disorders may be the consequence of derangements of the metabolism of the lipoprotein proteins. More information is presently being gathered on the mechanism of formation of these lipoproteins and on the factors regulating this process. It is also hoped that some information will be gained on the mechanism of biological degradation of these macromolecules, which is now essentially unknown. In the meantime, it has been fairly well established that increased VLDL production and subsequent elevation in plasma may be dependent on carbohydrate intake, thus providing a useful rationale for the treatment of this hyperlipoproteinemia (type IV). A similar carbohydrate dependence has been shown for type III hyperlipoproteinemia and this has again provided a good basis for its treatment. On the other hand, the elevation of serum β lipoproteins characteristic of type II hyperlipoproteinemia has not yet been explained, and the limited value of dietary treatment has also been recognized. It is thus evident, even from this superficial analysis, that a study of the lipoprotein abnormalities in terms of type of deficiency and mechanism of onset has both theoretical and practical implications. In the latter respect one of the most important problems is that of early recognition of these disorders and prevention of possible complications. It is established, for instance, that hyperlipoproteinemias, whether primary or secondary (nephrotic syndrome, hypothyroidism, diabetes mellitus, etc.) may favor atherosclerosis.

Expected Developments

Recent advances in the chemistry of the polypeptide moieties of the various human plasma lipoproteins have revealed evidence of structural complexities unsuspected at the time of the classification proposed by Fredrickson and Lees.[3] The opinion was then prevalent that there was one single type of HDL protein or apo A, one single protein in LDL or apo B, and that VLDL was probably made up of a mixture of apo A and apo B. It appears now, however, that VLDL has at least two additional polypeptide chains unrelated to A and B,[1, 5] and that the A itself has more than one component.[14, 16, 17, 19, 20] As these new units are characterized in detail, and further heterogeneities are possibly detected in the other lipoprotein fractions, a need for a more detailed classification of dyslipoproteinemia may become necessary. This, in turn, may call for the development of new techniques for the identification of these new units, since such techniques are likely to be more sophisticated than the ones currently in use. One of the problems will be that of adapting the techniques to the facilities of an ordinary clinical chemistry laboratory.

REFERENCES

1. Brown, W. V., Levy, R. I., and Fredrickson, D. S.: Studies on the proteins of the very low density lipoproteins. Fed. Proc., *28*:666, 1969.
2. Charman, R. C., and Landowne, R. A.: Separation of human plasma lipoprotein by electrophoresis on cellulose acetate. Analyt. Biochem., *19*:177, 1967.

3. Fredrickson, D. S., and Lees, R. S.: System for phenotyping hyperlipoproteinemia. Circulation, *31*:321, 1965.
4. Fredrickson, D. S., Levy, R. I., and Lees, R. S.: Fat transport in lipoproteins—An integrated approach to mechanisms and disorders. New Eng. J. Med., *276*:34, 94, 148, 215, 273, 1967.
5. Gustafson, A., Alaupovic, P., and Furman, R. H.: Studies of the composition and structure of serum lipoproteins. Separation and characterization of phospholipid-protein residues obtained by partial delipidization of very low density lipoproteins of human serum. Biochemistry, *5*:632, 1966.
6. Hatch, F. T., and Lees, R. S.: Practical methods for plasma lipoprotein analysis. Adv. Lipid Res., *6*:1, 1968.
7. Lees, R. S., and Hatch, F. T.: Sharper separation of lipoprotein species by paper electrophoresis in albumin-containing buffer. J. Lab. Clin. Med., *61*:518, 1963.
8. Lewis, L. A.: Screening of serum lipoprotein abnormalities: Comparison of ultracentrifugal paper and thin-layer starch-gel electrophoresis techniques. Lipids, *4*:60, 1969.
9. Narayan, K. A.: The influence of polyacrylamide concentration on human and rat serum lipoprotein patterns. Analyt. Biochem., *18*:582, 1967.
10. Nichols, A. V.: Human serum lipoproteins and their interrelationship. Adv. Biol. Med. Phys., *11*:109, 1967.
11. Noble, R. P.: Electrophoretic separation of plasma lipoproteins in agarose gel. J. Lipid Res., *9*:693, 1968.
12. Noble, R. P., Hatch, F. T., Mazrima, J. A., Lindgren, F. T., Jensen, L. C., and Adamson, G. L.: Comparison of lipoprotein analysis by agarose gel and paper electrophoresis with analytical ultracentrifugation. Lipids, *4*:55, 1969.
13. Norum, K. P., and Gjone, E.: Familial plasma lecithin:cholesterol acyltransferase deficiency. Biochemical study of a new inborn error of metabolism. Scand. J. Clin. Lab. Invest., *20*:231, 1967.
14. Rudman, D., Garcia, L. A., Abell, L. L., and Akgun, S.: Observations on the protein components of human plasma high- and low-density lipoproteins. Biochemistry, *7*:3136, 1968.
15. Scanu, A.: Factors affecting lipoprotein metabolism. Adv. Lipid Res., *3*:63, 1965.
16. Scanu, A., Reader, W., and Edelstein, C.: Molecular weight and subunit structure of human serum high density lipoprotein after chemical modification by succinic anhydride. Biochim. Biophys. Acta, *160*:32, 1968.
17. Scanu, A., Toth, J., Edelstein, C., Koga, S., and Stiller, E.: Fractionation of human serum high density lipoprotein in urea solutions. Biochemistry, *8*:3309, 1969.
18. Schumaker, V. W., and Adams, G. H.: Circulating lipoproteins. Ann. Rev. Biochem., *38*:113, 1969.
19. Shore, B., and Shore, V.: Heterogeneity in protein subunits of human serum high density lipoprotein. Biochemistry, *7*:2773, 1968.
20. Shore, V., and Shore, B.: Some physical and chemical studies of two polypeptide components of high density lipoproteins of human serum. Biochemistry, *7*:3396, 1968.

950 East 59th Street
Chicago, Illinois 60637

IV. THE ACTION OF HORMONES

Introduction

When 10 million women in the United States are made infertile by taking a "pill," the usefulness of hormone therapy for the social community cannot be questioned. Another question, however, should be raised: Is present knowledge of the site of production and the locus of the action of hormones advanced in enough detail to permit an unprecedented wholesale interference with biological processes with general effects small enough to justify a needed check on population growth? The immediate problem is considered in the first paper of this section. The recognition and function of a hormone and the consequences of its hyperfunction or hypofunction are presented in discussions on two "new" hormone entities: erythropoietin—a particular province of Chicago medicine—and glucagon—of renewed interest because of its stimulating effect on cardiac muscle.

Two papers of the section deal with cellular structure and function of hormones as illustrated by insulin. The last two decades have witnessed a remarkable concern of medical investigators in the study of cellular function and subcellular structures with relation to adjustments and maladjustments of organic function. This constitutes an attempt to establish a truly molecular basis for many problems in medicine. The biochemical isolation of proinsulin, and the role of insulin in protein synthesis as one example of the mechanisms of hormone action, are given as examples of the new era of molecular medicine.

H. H. H.

Current Status of Oral Contraceptives

ROBERT E. CLEARY, M.D.*

RAJAI M. DAJANI, M.D.**

Oral contraceptives are among the very few drugs that have achieved rapid and widespread acceptance. In 1967 over 7 million women in the United States were using "the pill".[45] The virtual 100 per cent effectiveness of this preparation as well as the intense interest in a method of combating the spiraling population growth were prime factors in its rapid acceptance.[14, 25]

This article will have as its prime objective a review of certain aspects of the untoward systemic effects and risks cited in the recent literature.

PREPARATIONS

The basic composition of "the pill" is a combination of a synthetic progestin and an estrogen in different dosages. The synthetic estrogens in use are mestranol and ethinyl estradiol. The progestins employed are listed in Table 1.

Oral contraceptives fall into three major regimens: (1) combined, (2) sequential, and (3) continuous progestin. In the combined regimen a

Table 1. *Progestins Employed in Oral Contraceptives*

C_{19} Norsteroid derivatives
Norethynodrel
Norethindrone
Ethynodiol diacetate
Norgestrel
Dimethisterone
17-Hydroxyprogesterone derivatives
Medroxyprogesterone acetate
Chlormadinone

*Assistant Professor, Department of Obstetrics and Gynecology, University of Chicago, Chicago Lying-In Hospital, Chicago, Illinois

**Instructor, Department of Obstetrics and Gynecology, University of Chicago, Chicago Lying-In Hospital, Chicago, Illinois

Table 2. *Currently Available Combined Oral Contraceptives*

TRADE NAME	PROGESTIN AND DOSE mg. per day	ESTROGEN AND DOSE mg. per day
Enovid 5	Norethynodrel 5	Mestranol 0.075
Enovid E	Norethynodrel 2.5	Mestranol 0.10
Ovulen-21	Ethynodiol diacetate 1.0	Mestranol 0.10
Ovulen-28*		
Norinyl 2	Norethindrone 2.0	Mestranol 0.10
Norinyl-1 (Noriday*)	Norethindrone 1.0	Mestranol 0.05
Norinyl 1/80	Norethindrone 1.0	Mestranol 0.08
Ortho Novum 10	Norethindrone 10.0	Mestranol 0.06
Ortho Novum 2	Norethindrone 2.0	Mestranol 0.10
Ortho Novum 1	Norethindrone 1.0	Mestranol 0.05
Ortho Novum 1/80	Norethindrone 1.0	Mestranol 0.08
Provest	Medroxyprogesterone acetate 10.0	Ethinyl estradiol 0.05
Ovral	Norgestrel 0.5	Ethinyl estradiol 0.05
Norlestrin 2.5	Norethindrone 2.5	Ethinyl estradiol 0.05
Norlestrin 1 mg.	Norethindrone 1.0	Ethinyl estradiol 0.05
Norlestrin 28*		
Norlestrin Fe†		

*Days 22 to 28, seven inert tablets.
†Days 22 to 28, iron supplement tablets.

Table 3. *Currently Available Sequential Oral Contraceptives*

TRADE NAME	PROGESTIN AND DOSE mg. per day 21 to 25	ESTROGEN AND DOSE mg. per day 5 to 25
Oracon	Dimethestrone 25	Ethinyl estradiol 0.10
C-Quens	Chlormadinone 2.0	Mestranol 0.08
Ortho Novum SQ	Norethindrone 2.0	Mestranol 0.08
Norquen	Norethindrone 2.0	Mestranol 0.08

progestin and an estrogen combined in a tablet are administered daily for 20 or 21 days. In the sequential regimen an estrogen is administered for 14 to 16 days, followed by the daily administration of the combined estrogen and progestin tablet for 5 or 6 days. Tables 2 and 3 list the composition and the dosages of the currently available medications.[2, 15, 31]

Currently under investigation is the continuous daily administration of a progestin, chlormadinone, in a dose of 0.5 mg. per day, or norqestrel, in a dose of 0.05 mg. per day.[20, 43] These preparations have been found effective, but their widespread acceptance may be delayed by the higher incidence of bleeding irregularities encountered with their use.[20, 43]

MODE OF ACTION

Oral contraceptives may act to a varying degree at each of the following areas: (1) hypothalamus and pituitary, (2) endometrium, (3) cervical mucus, and (4) ovum and sperm metabolism.[11] Despite major

advances in methodology to study these areas, the mechanisms of action by which oral contraceptives inhibit fertility at each site is not completely known. The hypothesis that estrogen or progestin can inhibit the release of gonadotropins was formulated originally from data on urinary gonadotropins.[22, 42, 53, 65, 66, 69, 77] These findings were reinvestigated with the development of the radioimmunoassay for plasma FSH and LH.[36] Several investigators have shown that the combined preparations inhibit the early FSH elevation and abolish the midcyclic LH and FSH peaks.[6, 36, 68] The sequential preparations interfere with the cyclic secretions of FSH and LH.[36, 68]

Changes in plasma levels of FSH and LH do not necessarily mean that the pituitary is the site of action. In fact, the actual site of action may be the hypothalamus or a higher brain center.[59]

Oral contraceptives produce definite changes in the endometrial glands and stroma that could prevent implantation of the blastocyst. The most marked changes are produced by the combined preparations, which effect regression of the endometrial gland early in the cycle followed by secretory exhaustion in the premenstrual phase.[54]

Aside from the endometrium, another target area of the reproductive tract affected by oral contraceptives is cervical mucus. With combined therapy or the continuous low dose progestin, certain definitive changes are produced in cervical mucus. The pattern of "fern" present during the proliferative phase disappears, the volume decreases, and the cellular content and the viscosity increase.[5, 47] By far the most important change induced by the progestin is the inhibition of sperm migration in cervical mucus.[43, 47]

The exact role of estrogen or the progestin in altering capacitation of sperm or the metabolism of the blastocyst in the human reproductive tract remains to be determined. Capacitation of spermatozoa in the reproductive tract of a rabbit can be inhibited by progesterone.[7] It is

Table 4. *Comparative Pregnancy Rates Among Different Contraceptive Methods*

METHOD	PREGNANCY RATE PER 100 WOMAN YEARS		
	TIETZE 1959	VENNING 1961	GARCIA 1968
Douche	–	31	37.8
Foam tablets	13–49	–	22
Suppositories	8–27	–	–
Jelly	11–38	20	20
Coitus interruptus	–	18	16
Rhythm (safe period)	14	24	14.0
Condom	11	14	14.9
Diaphragm	6–29	12	12
Cervical cap	8	–	–
IUD	–	–	3.9
Sequential orals	–	–	1.4
Combined orals	–	1.2	0.2

therefore possible that the low dose progestin may produce an additional antifertility effect by its action at this site.

EFFECTIVENESS

The combined "pill" is currently the most effective form of contraception. The sequentials evidently have a higher pregnancy rate, but this is still far below the other available contraceptive means (Table 4).[25, 26, 29, 44, 70, 73]

SIDE EFFECTS

Carbohydrate Metabolism

The literature is now replete with reports of impaired glucose tolerance in women taking oral contraceptives.[27, 55, 63, 64, 79, 81] One recent article attributes the increased incidence of impaired glucose tolerance in previous studies to the inclusion of obese and prediabetic individuals as study subjects.[81]

Further studies into the cause of impaired glucose tolerance have revealed the presence of hyperinsulinism, a very definite increase in human growth hormone, and a relative hypertriglyceridemia.[30, 62, 63, 80, 81]

One hypothesis is that estrogens enhance pituitary release of growth hormone.[23] Growth hormone then induces a peripheral antagonism to insulin.[10] So long as the compensating increase in insulin secretion can be maintained, the glucose tolerance test may be normal.[81]

Effects on Liver Function

Jaundice associated with the use of oral contraceptives has been reported.[39, 49, 52] Ockner et al. described 40 such cases in which liver biopsies revealed canalicular and hepatocellular bile stasis.[49] These changes were completely reversible when the drug was withdrawn.[29, 49]

There are many reports in the literature referring to various alterations in liver function tests.[1, 3, 19, 26, 28, 39, 41, 52, 72] The most consistently noted finding is an increase in Bromsulfalein retention,[19, 26, 39, 50, 58] which is apparently dose related.[1, 50, 58] There is no consistent agreement among authors concerning the changes in serum bilirubin, SGOT, SGPT, thymol turbidity, cephalin flocculation, or alkaline phosphatase.[1, 3, 19, 50, 57, 58, 71] In view of the lack of evidence of permanent histological or functional damage to the liver, the term "liver dysfunction" would be more appropriate than liver damage.[50, 71]

Virtually all authors agree that the use of oral contraceptives is to be avoided in patients with a previous history of cholestatic jaundice of pregnancy, primary biliary cirrhosis, or constitutional hyperbilirubinemia.[1, 14, 50, 58] Since the changes described are dose related, the use of the lower dosage medication is also advisable.

Thyroid Function

Studies on free thyroxine levels and absolute thyroxine turnover have established the normal thyroid status of those taking oral contra-

Table 5. *Estrogen-Induced Alterations in Thyroid Function Studies*

THYROID FUNCTION STUDIES	ALTERATION
PBI	Increased
^{131}I-T_3 red cell or resin	Decreased
Free thyroxine index	None
RAI uptake	None
Free thyroxine	None
Total thyroxine	Increased

ceptives.[28] Estrogens induce a definite increase in thyroxine-binding globulin.[51] The result is an elevation of the serum PBI and a decrease in the uptake of labeled triiodothyronine by red blood cells or by a resin.[28, 51] The predictable inverse changes in these two tests allow calculation of the free thyroxine index. This is a suitable parameter to evaluate thyroid function when patients are taking estrogens.[8] Other alterations in thyroid function studies are listed in Table 5.[8, 28, 46, 51]

Amenorrhea

A number of recent reports have focused attention on the occurrence of amenorrhea after discontinuing the "pill".[12, 56, 61, 78] In some of the earlier cases cited, amenorrhea persisted for only a few months. The cases of short duration must be distinguished from absence of withdrawal flow that can occur with the combined pill. Shearman reported 22 cases in which amenorrhea was present for at least 12 months. Of interest was the presence of pre-existing menstrual abnormalities in eight of these patients.

The onset of amenorrhea after discontinuing oral contraceptives does not necessarily prove a cause and effect relationship. The cause of secondary amenorrhea may have existed prior to the initiation of oral contraceptive therapy. The possibility of post-"pill" amenorrhea must be considered carefully before placing patients with menstrual irregularities on oral contraceptives.

Effects on Blood Coagulation

Since the possibility of an association between thromboembolic disease and the use of oral contraceptives was first suspected, considerable attention and research activity has been focused on coagulation mechanisms in women taking the "pill".[31, 33, 37] These studies attempt to link alterations of clotting factors produced by oral contraceptives with the pathophysiology of venous thrombosis and thromboembolic disease. Most of the studies demonstrate an increase in factors II, VII, and X, and factor VII-X (P and P) complex, with the use of oral contraceptives.[9, 14, 18, 32, 34, 48, 57] There is no decrease in fibrinolytic activity as seen in pregnancy.[32, 48] Donayre et al. have even noted an increase in fibrinolytic activity.[13] No appreciable change in platelet count or adhesiveness has been reported.[4, 32, 57] Most authors agree that these alterations in coagulation factors do not predispose to venous thrombosis.[9, 32, 34, 48, 57, 67]

Table 6. *Comparative Mortality and Morbidity Figures from Venous Thromboembolic Disease and Pregnancy Among Women of Ages 20 to 44 in Britain per 100,000 Woman Years**

CATEGORY	MORTALITY	MORBIDITY (Hospitalization)
Users of oral contraceptives	2.7	47
Nonusers of oral contraceptives	0.35	5
Pregnant	40.2	–

*Data modified from Inman and Vessey, 1968, and Vessey and Doll, 1968.

Thromboembolism

Recent evidence suggests a statistical correlation between deep venous thrombosis and pulmonary embolism and the use of oral contraceptives.[33, 37, 38] The Medical Research Council of Britain found an increased morbidity and increased mortality among users of oral contraceptives (Table 6).[35, 74, 76] Morbidity was defined as the number of patients requiring hospitalization for venous thromboembolic disease.

A cause and effect relationship has not been established.[16, 17, 21, 24, 45, 60] This question will remain debatable until more information can be obtained from the four ongoing prospective studies in England and the United States.[74] It is interesting to note in this respect that the incidence of spontaneously occurring thromboembolic disease seems to be higher in Europe than in the United States.[14] Vessey and Weatherall have cautioned against the projection of the British figures to any other country because of the possible difference in incidence of spontaneously occurring thromboembolic disease.[76]

Not included among the British group labeled "non-users of oral contraceptives" are those who would have become pregnant using other less effective forms of contraception.[45] These women would then be subject to the increased morbidity of venous thrombosis and thromboembolic disease and the increased risk of mortality from these and other complications of pregnancy.

CONCLUSION

When 7 million women are taking oral contraceptives, the relative hazards and safety of the "pill" must be considered carefully. There is no doubt that the ideal antifertility agent should be completely safe and limit its action to the reproductive tract. Until this is achieved, the physician must choose between the overall hazards of pregnancy when other less effective contraceptives fail and the risk of the "pill".

Newer developments in the area of the low dose progestin, immunological methods for conception control, or nonsteroidal antifertility agents, may further reduce the risk of contraceptive therapy.[40, 72]

REFERENCES

1. Allan, J. S.: Hormone contraception; liver function and other biochemical studies. *In* Advanced Concepts in Contraception. New York, Excerpta Medica, 1968, p. 14.
2. Behrman, S. J.: Which pill to choose. Hosp. Practice, *4*:34, 1969.
3. Besch, P. K., Vorys, N., Villery, J. L., Stevens, V., and Barry, R.: Some systemic effects of a progestational agent. Metabolism, *14*:387, 1965.
4. Bolton, C. H., Hampton, J. R., and Mitchell, J. R. A.: Effects of oral contraceptive agents on platelets and plasma phospholipids. Lancet, *1*:1336, 1968.
5. Bowman, J. A.: The effect of norethindrone-mestranol on cervical mucus. Amer. J. Obstet. Gynec., *102*:1039, 1968.
6. Cargille, C. M., and Ross, G. T.: Oral contraceptives and follicle stimulating hormone. Lancet, *1*:924, 1968.
7. Chang, M. C.: Capacitation of rabbit spermatozoa in the uterus. Endocrinol., *63*:619, 1958.
8. Clark, F., and Horn, D. B.: Assessment of thyroid function by the combined use of the serum protein-bound iodine and resin uptake of ^{131}I-triiodothyronine. J. Clin. Endocrinol., *25*:39, 1965.
9. Connell, E., Goldstein, R., and Reiner, S.: Blood coagulation and oral contraceptives. *In* Advances in Planned Parenthood. New York, Excerpta Medica, 1968, Vol. 4, p. 117.
10. Daughaday, W. H., and Kipnis, D. M.: The growth-promoting and anti-insulin actions of somatotropin. Recent Progr. Hormone Res., *22*:49, 1966.
11. Diczfalusy, E.: Mode of action of contraceptive drugs. Amer. J. Obstet. Gynec., *100*:136, 1968.
12. Dodek, O. I., and Kotz, H. I.: Syndrome of anovulation following the oral contraceptives. Amer. J. Obstet. Gynec., *98*:1065, 1968.
13. Donayre, J., and Pincus, G.: Effect of Enovid on blood clotting factors. Metabolism, *14*: 418, 1965.
14. Drill, V. A.: Oral Contraceptives. New York, McGraw-Hill Book Co., 1966.
15. Drill, V. A.: Endocrine properties of oral contraceptives. Metabolism, *14*:295, 1965.
16. Drill, V. A., and Calhoun, D.: Oral contraceptives and thromboembolic disease. J.A.M.A., *206*:77, 1968.
17. Drill, V. A., and Calhoun, D.: Letter to the editor. J.A.M.A., *207*:1151, 1969.
18. Egeberg, O., and Owren, P. A.: Oral contraception and blood coagulation. Brit Med. J., *1*:220, 1963.
19. Eisalo, A., Jarvinen, P. A., and Luukainen, T.: Liver function tests during intake of contraceptive tablets in premenopausal women. Brit. Med. J., *1*:220, 1963.
20. Foss, G. L., Svendsen, E. K., Fotherby, K., and Richard, D. J.: Contraceptive action of continuous low dose of norgestrel. Brit. Med. J., *4*:489, 1968.
21. Frank, R., and Tietze, C.: Oral contraceptives and phlebitis. Amer. J. Obstet. Gynec., *93*:122, 1965.
22. Flowers, C., Stevens, V. C., Vorys, N., Miller, A. T., and Jensen, L.: The effects of suppression of menstruation with ethynodiol diacetate upon the pituitary ovary and endometrium. Amer. J. Obstet. Gynec., *96*:784, 1966.
23. Frantz, A. B., and Rabkin, M. T.: Effects of estrogen and sex difference on secretion of human growth hormone. J. Clin. Endocrinol., *25*:1470, 1965.
24. Fredriksen, H., and Ravenholt, R. T.: Thromboembolic disorders and oral contraceptives. J.A.M.A., *207*:1150, 1969.
25. Garcia, C.: The oral contraceptive. Amer. J. Med. Sci., *253*:114, 1967.
26. Garcia, C. R.: Medical and metabolic effects of oral contraceptives and their implications. Clin. Obstet. Gynec., *11*:669, 1968.
27. Gershberg, H., Javier, Z., and Hulse, M.: Glucose tolerance in women receiving an ovulatory suppressant. Diabetes, *13*:378, 1964.
28. Goolden, A. W. G., Gartside, J., and Sanderson, C.: Thyroid status in pregnancy and in women taking oral contraceptives. Lancet, *1*:12, 1967.
29. Guttmacher, A.: Contraceptive problems. Postgrad. Med., *41*:233, 1967.
30. Hazzard, W. R., Spiger, M. J., Bagdade, J. D., and Bierman, E. L.: Studies on the mechanism of increased plasma triglyceride levels. New Eng. J. Med., *280*:471, 1969.
31. Hertz, R., and Bailar, J.: Estrogen-progestogen combinations for contraception. J.A.M.A., *198*:1000, 1966.
32. Hilden, M., Amris, C. J., and Starup, J.: Hemostatic mechanisms in oral contraception. Acta Obstet. Gynec. Scandinav., *46*:562, 1967.
33. Hobel, C. J., and Mishell, C. R.: Pulmonary embolism and oral steroidal contraceptives. Amer. J. Obstet. Gynec., *101*:994, 1968.
34. Hougie, C., Rutherford, R., Banks, A. L., and Coburn, W. A.: Effects of progestin-estrogen oral contraceptives on blood clotting factors. Metabolism, *14*:411, 1965.
35. Inman, W. H. W., and Vessey, M. P.: Investigation of death from pulmonary, coronary and cerebral thrombosis in women of childbearing age. Brit. Med. J., *2*:193, 1968.

36. Jaffe, R. B., and Midgley, A. R., Jr.: Current status of human gonadotropin radioimmunoassay. Obstet. Gynec. Survey, *24*:200, 1969.
37. Lancet: Oral contraceptives and thromboembolism (editorial). *1*:1039, 1967.
38. Lancet: Oral contraceptives and thromboembolism (editorial). *1*:962, 1968.
39. Larson-Cohn, U.: Oral contraceptives and liver function tests. Brit. Med. J., *1*:1414, 1965.
40. Laurence, K. A.: Immunological conception control. *In* Advanced Concepts in Contraception. New York, Excerpta Medica, 1968, p. 102.
41. Linthorst, G.: Liver function after long term progestational treatment with or without estrogens. Brit. Med. J., 2:920, 1964.
42. Loraine, J. A., Bell, E. T., Harkness, R. A., Mears, E., and Jackson, M. C. N.: Hormone excretion patterns during and after the long term administration of oral contraceptives. Acta Endocrinol., *50*:15, 1965.
43. Martinez-Manatou, J., Velasquez-Giner, J., Cortes-Gallegos, V., Aznar, R., Rojas, B., Guitterez-Najar, A., and Rudel, H. W.: Daily progestogen for contraception. Brit. Med. J., 2:730, 1967.
44. McQuarrie, H. G.: Oral contraceptives. MED. CLIN. N. AMER., *51*:1261, 1967.
45. Meeker, C. I.: Use of drugs and intrauterine devices for birth control. New Eng. J. Med., *280*:1058, 1969.
46. Mishell, D. R., Jr., Colodny, S. Z., and Swanson, K. A.: The effect of an oral contraceptive on tests of thyroid function. Fertil. Steril., *20*:335, 1969.
47. Moghissi, K. S.: Cyclic changes of cervical mucus in normal and progestin-treated women. Fertil. Steril., *17*:663, 1966.
48. Nilsson, I. M., and Kullander, S.: Coagulation and fibrinolytic studies during use of gestogens. Acta Obstet. Gynec. Scandinav., *46*:286, 1967.
49. Ockner, R., and Davidson, C. S.: Hepatic effects of oral contraceptives. New Eng. J. Med., *276*:331, 1967.
50. Okamura, H.: Effects of various oral gestogens on the liver function. J. Jap. Obstet. Gynec. Soc., *15*:117, 1968.
51. Oppenheimer, J. H.: Plasma proteins in distribution and metabolism of thyroid hormones. New Eng. J. Med., *278*:1153, 1968.
52. Orellana-Alcalde, J. M., and Dominguez, J. P.: Jaundice and oral contraceptive drugs. Lancet, 2:1278, 1966.
53. Ostergaard, E., Arends, J., Hamburger, C., and Johnsen, S. G.: Postponement of menses, suppression of urinary gonadotropins, 17-KGS and 7-KS in normal women treated with megesterol acetate/mestranol. Acta Endocrinol., *53*:13, 1966.
54. Pincus, G., Rock, J., and Garcia, C.: Effects of certain 19-norsteroids upon reproductive processes. Ann. N.Y. Acad. Sci., *71*:677, 1958.
55. Posner, N. A., Silverstone, F. A., Pomerance, W., and Singer, N.: Oral contraceptives and intravenous glucose tolerance. Obstet. Gynec., *29*:87, 1967.
56. Rankin, R. P.: Prolonged anovulation subsequent to oral progestins. Amer. J. Obstet. Gynec., *103*:919, 1969.
57. Robinson, R.: Effect of estrogen-progestin on clotting factors. Amer. J. Obstet. Gynec., *99*:163, 1967.
58. Schaffner, F.: Effect of oral contraceptives on liver. J.A.M.A., *198*:1019, 1966.
59. Schally, A. W., Carter, W. H., Saito, M., Arimura, A., and Bowers, C. Y.: Studies on site of action of oral contraceptive steroids. J. Clin. Endocrinol., *28*:1747, 1968.
60. Seltzer, C.: Letter to the editor. J.A.M.A., *207*:1152, 1969.
61. Shearman, R. P.: Investigation and treatment of amenorrhea developing after treatment with oral contraceptives. Lancet, *1*:325, 1968.
62. Spellacy, W. N., Buhi, W. C., and Bendel, R. P.: Growth hormone alterations by a sequential-type oral contraceptive. Obstet. Gynec., *33*:506, 1969.
63. Spellacy, W. N., and Carlson, K. L.: Plasma insulin and blood glucose levels in patients taking oral contraceptives. Amer. J. Obstet. Gynec., *95*:474, 1966.
64. Spellacy, W. N., Carlson, K. L., and Birk, S. A.: Carbohydrate metabolic studies after six cycles of combined type oral contraceptive tablets. Diabetes, *16*:590, 1967.
65. Stevens, V. C., Goldzieher, J. W., and Vorys, N.: Effect of mestranol and chlormadinone acetate on urinary excretion of FSH and LH. Amer. J. Obstet. Gynec., *102*:95, 1968.
66. Stevens, V. C., and Vorys, N.: The regulation of pituitary function by sex steroids. Obstet. Gynec. Surv., 22:781, 1967.
67. Swyer, G. I. M., and Little, V.: Abscence of hepatic impairment in long-term contraceptive users. Brit. Med. J., *1*:1412, 1965.
68. Swerdloff, R. A., and Odell, W. D.: Serum luteinizing and follicle stimulating hormone levels during sequential and nonsequential contraceptive treatment of eugonadal women. J. Clin. Endocrinol., 29:157, 1969.
69. Taymor, M. L.: Effect of synthetic progestins on pituitary gondatrophin secretion. J. Clin. Endocrinol., 24:803, 1964.
70. Tietze, C.: Clinical effectiveness of contraceptive methods. Amer. J. Obstet. Gynec., *78*:65, 1959.

71. Tyler, E. T.: Oral contraceptives. Brit. Med. J., 2:843, 1964.
72. Tyler, E. T.: Antifertility agents. Ann. Rev. Pharmacol., 7:381, 1967.
73. Venning, G. P.: The contraceptive pill. Brit. Med. J., 2:899, 1961.
74. Vessey, M. P.: Oral contraceptives and thromboembolic disease. Amer. Heart J., 77:153, 1969.
75. Vessey, M. P., and Doll, R.: Investigation of relation between use of oral contraceptives and thromboembolic disease. Brit. Med. J., 2:199, 1968.
76. Vessey, M. P., and Weatherall, J.: Venous thromboembolic disease and use of oral contraceptives. Lancet, 2:94, 1968.
77. Vorys, N., Villery, J., and Stevens, V.: The effects of sex steroids on gonadotropins. Amer. J. Obstet. Gynec., 93:641, 1965.
78. Whitelaw, M. J., Nola, V. F., and Kalman, C. F.: Irregular menses, amenorrhea and infertility following synthetic progestational agents. J.A.M.A., 195:780, 1966.
79. Wynn, V., and Doar, J. W. H.: Some effects of oral contraceptives on carbohydrate metabolism. Lancet, 2:715, 1966.
80. Wynn, V., and Doar, J. W. H.: Some effects of oral contraceptives on serum-lipid and lipoprotein levels. Lancet, 2:720, 1966.
81. Yen, S. S. C., and Vela, P.: Effects of contraceptive steroids on carbohydrate metabolism. J. Clin. Endocrinol., 28:1564, 1968.

Department of Obstetrics and Gynecology
University of Chicago
Chicago, Illinois 60637

Current Status of Erythropoietin

SANFORD B. KRANTZ, M.D.*

Erythropoietin is a hormone that acts on the bone marrow to increase red cell production while not affecting white cell and platelet production.[36] It is found in the plasma and urine of many mammalian species, as well as in birds and fish, after they are made hypoxic,[36] and is also found in the urine of normal individuals.[1, 2, 10] Its presence has been demonstrated in concentrated extracts of normal plasma,[18, 40] but it generally cannot be detected in untreated normal plasma, presumably because of the relative insensitivity of the bioassay.

Production of the hormone is increased under conditions that lead to tissue hypoxia and is decreased by conditions that create tissue hyperoxia.[30] Thus atmospheric hypoxia and anemic hypoxia lead to increased erythropoietin production and increased erythropoiesis, while atmospheric hyperoxia, or plethora due to hypertransfusion, lead to decreased erythropoietin and decreased erythropoiesis. Similarly, thyroid hormone or dinitrophenol, which increase oxygen consumption, also increase erythropoietin production, while hypophysectomy, starvation, or hypothyroidism, which reduce oxygen consumption, decrease erythropoietin production. A change in the level of tissue oxygenation is thus followed by an alteration in erythropoietin production and a compensatory modification of red cell production which tends to restore the previous level of tissue oxygenation.[30]

This fact has been utilized in the design of the current assay for the hormone. Mice are either hypertransfused with isologous red cells[31] or are maintained in hypoxic chambers for several weeks before being brought into a normal atmosphere.[5] In the former method an immediate plethora is produced which results in hyperoxia and reduced erythropoiesis. In the latter method the hypoxia stimulates erythropoietin production and erythropoiesis, so that the mice produce their own plethora. When the animals are then placed in a normal atmosphere the red cell mass is increased in relation to the physiological need and erythropoiesis ceases, as it does in the hypertransfused mice. In this

*Staff Associate, Clinical Pathology Department, National Institutes of Health, Bethesda, Maryland; Assistant Professor of Medicine (Hematology), Department of Medicine, University of Chicago, Chicago, Illinois

condition the animals are much more sensitive to the administration of erythropoietin, which increases the reticulocyte count or the incorporation of radioiron into new red cells. The percentage of injected radioiron which is found in the red cells of the peripheral blood is generally used to measure the effect of the hormone, since it is proportional to the amount of hormone administered.[6]

The principal organ controlling the production of erythropoietin is the kidney. This has been shown by many experiments in which nephrectomy was followed by a rapid decline of blood erythropoietin levels.[7, 29, 49] Erythropoietin has also been demonstrated in extracts of renal tissue.[27, 41] The incubation of light mitochondrial extracts of the kidney with normal plasma appears to increase the amount of erythropoietin, suggesting that the kidney may produce erythropoietin by an enzymatic action on a plasma substrate similar to the way in which renin produces angiotensin.[20] The enzyme and substrate have not been isolated, however, and it is still possible that the plasma serves to stabilize the product of an enzymatic reaction. A small amount of erythropoietin can be detected in anephric individuals,[9, 47, 50] and this extrarenal production appears to be enhanced by extreme degrees of tissue hypoxia.[16, 50] The site of extrarenal erythropoietin is still unknown.

Erythropoietin has been purified by a factor of 930,000 to 8,000 units per mg. of protein, but is still not entirely free of impurities.[19] It appears to be a glycoprotein with a molecular weight of 60,000 to 70,000.[19] The main source of the hormone for purification procedures is the plasma of anemic animals[19] or the urine of anemic human beings.[19, 43] A principal hindrance to the isolation of erythropoietin has been the instability of highly purified hormone preparations[19] compared to the relative stability of crude preparations.[34] An antibody has been made which neutralizes the biological activity of human urinary erythropoietin and also the activity of plasma erythropoietin from several animal species.[56] This, together with several common physical-chemical characteristics, and a similar biological action in the polycythemic mouse assay, suggests that erythropoietins from different mammalian sources are very similar.

The hormone acts on an erythropoietin-responsive cell (ERC) that precedes the morphologically well-defined erythroblasts to initiate their differentiation into erythroblasts.[24, 31] These erythroblasts then mature into reticulocytes and red blood cells without further need for additional erythropoietin.[8, 56] There is some evidence, however, to indicate that large amounts of the hormone may play an additional role in accelerating erythroblast maturation and the release of reticulocytes into the blood.[58] The ERC is a descendant of the so-called colony-forming cell, or else it is a colony-forming cell in a particular state of its cell cycle. These colony-forming cells are pluripotential stem cells which, when transfused into irradiated hosts, form erythrocytic, granulocytic, or megakaryocytic clones in the spleen and marrow of the hosts.[42, 59] After many years of study the morphological identity of these stem cells is still undetermined, but the lymphocyte, transitional cell, and monocytoid cell are the current favorites.

The earliest effect of erythropoietin on marrow cells is an increase in RNA synthesis.[38] This RNA is made up of several components, which appear to include ribosomal precursor RNA, transfer RNA, and messenger RNA.[21] It is believed that the hormone acts on morphologically undeveloped cells to allow particular areas of the DNA to become active and express a distinct genetic capacity. These genes produce a distinct messenger RNA, which in turn produces the proteins such as hemoglobin that characterize the newly formed erythroblast. A very early increase in DNA synthesis has also been noted after the application of the hormone to erythropoietic tissue.[54] This is soon followed by an increase in iron incorporation into the cells,[28] an increase in the manufacture of the cell stroma, and an increase in hemoglobin synthesis.[37]

Erythropoietin is also actively involved in fetal erythropoiesis, although this process appears to be independent of the supply of *maternal* erythropoietin.[12, 13, 32] The hormone, however, has been demonstrated in the blood and urine of the human fetus by the 32nd week of gestation,[26] which suggests that the fetus may manufacture its own supply of the hormone. Increased erythropoietin levels also have been found in the amniotic fluid.[13] The amount of amniotic erythropoietin has been correlated with the degree of fetal hypoxia, indicating that it might be used as a parameter of fetal anemia.[11] The optical density of the amniotic fluid also correlates well with fetal anemia, however, and this is easier to measure.[11] Erythropoietin markedly stimulates erythropoiesis in the fetal mouse liver, but has no effect on yolk sac erythropoiesis.[4] Thus information about the action of erythropoietin is being extended to the study of embryonic and fetal erythropoiesis. The techniques developed for the investigation of erythropoietin in laboratory animals are also being applied to the study of clinical disorders of erythropoiesis.

RELATION OF ERYTHROPOIETIN TO CLINICAL DISEASE

Anemia

Anemias due to hemorrhage or hemolysis are generally accompanied by a rise in the blood and urinary levels of erythropoietin.[36] The increased amount of the hormone increases erythropoiesis in an attempt to compensate for the anemic hypoxia and re-establish the normal state. Anemias due to decreased marrow production of red cells may be divided into two classes: those associated with high erythropoietin levels and those associated with no increase in blood or urinary erythropoietin.[36] In the former case a primary marrow cell injury is generally present and the increased amount of erythropoietin, which results from the lack of adequate red cell production, is unable to restore normal erythropoiesis. In the latter case a lack of the hormone may be partly responsible for the decreased erythropoiesis.

The most evident anemia that may be due to decreased erythropoietin production is the anemia accompanying uremia. Most of the pa-

tients with this condition do not have elevated erythropoietin levels despite severe anemia.[22] Although this anemia may be partly due to other factors, such as a toxic depression of the marrow cells, or hemolysis and hemorrhage, the fact that uremic animals respond to the administration of the hormone[7, 29, 49] suggests that lack of erythropoietin is instrumental in producing the anemia and that replacement of the hormone might be of therapeutic benefit. In addition, renal transplantation has been followed by increased erythropoietin levels and increased erythropoiesis, in contrast to the observations made prior to the transplantation, which have shown that erythropoietin is generally not demonstrable unless the patients are severely anemic.[36] Additional studies have revealed that erythropoietin secretion can be clearly separated from the excretory capacity of the transplanted kidneys, since the two functions often return at different intervals after renal transplantation. Thus renal damage and not uremia per se is the primary cause of decreased erythropoietin production. Although there is a general correlation between increased blood urea nitrogen levels and decreased erythropoiesis, the two do not always correlate.[36]

The anemia associated with protein deprivation is probably due to reduced erythropoietin elaboration, since it is easily corrected in animals by small doses of erythropoietin.[17] In kwashiorkor, blood erythropoietin levels were found to be elevated.[44] The anemia was not corrected, however, unless hematinics were administered, which indicated that vitamin deficiencies probably produced a degree of primary marrow cell failure and a severe anemia in excess of that associated with the protein deficiency.[44] This presumably led to the increase in erythropoietin production.

In laboratory animals, chronic inflammation produced by turpentine abscesses resulted in an anemia that could be corrected by the administration of erythropoietin.[25] The anemias of patients with rheumatoid arthritis or chronic suppuration have responded to the administration of cobalt, which has been shown in laboratory animals to act through an increase in erythropoietin production.[36] Anemias in patients with neoplasms are generally associated with increased blood or urinary erythropoietin levels, although in some cases no increase was found.[36] This may be related to the patient's malnutrition.

A variety of endocrine deficiencies lead to anemia and this is thought to occur via the reduced elaboration of erythropoietin.[36] Castration of male animals leads to decreased erythropoiesis, which can be reversed by the administration of testosterone or erythropoietin. Further experiments have shown that treatment with androgens leads to an increase in blood or urinary erythropoietin levels, which apparently is one of the prime means by which the androgens affect erythropoiesis.[15, 46] Androgens have been shown to be of value in increasing erythropoiesis in patients with aplastic anemia, myelofibrosis, breast cancer, multiple myeloma, and several other neoplasms.[36] The administration of androgens enhances erythropoietin production and erythropoiesis in patients who already have high levels of the hormone.[48, 55] This result should not be surprising, since it may represent the difference between physiological and pharmacological effects of the hormone. This work suggests

that pharmacological doses of erythropoietin may prove to be useful in the future in the treatment of patients with a variety of anemias, even when a reduced capacity to produce the hormone is not the primary cause of the anemia.

Hypophysectomy also leads to an anemia that can be reversed by the administration of a variety of pituitary hormones or erythropoietin. ACTH, growth hormone, and TSH appear to increase erythropoiesis through an increase in erythropoietin production.[36] Thyroid hormone appears to act similarly, since the anemia of thyroidectomized animals does not respond to thyroid hormone after nephrectomy[57] and since antiserum to erythropoietin prevents an increase in erythropoiesis which otherwise follows triiodothyronine treatment.[14]

In pure red cell aplasia, a disease in which there is an isolated absence of erythroid precursors, erythropoietin levels are generally quite elevated.[36] Since a similar condition can be produced in mice by the administration of an anti-erythropoietin antibody, it is conceivable that an absence of the hormone could produce some cases of this disease. Recently one patient has been reported whose erythropoietin levels were not elevated despite a severe anemia.[33] An inhibitor to erythropoietin appeared to be present in this patient's plasma. In another patient with red cell aplasia, study of the marrow's response to erythropoietin in vitro led to the demonstration of an antibody to the erythroblasts and an inhibitor of heme synthesis in the patient's plasma.[39] Thus red cell aplasia may be due to an inhibitor to erythropoietin or to an inhibitor directed at the erythroblasts. In either case the development of the erythropoietin technology and its application to clinical disease has greatly helped to clarify the pathogenesis of this particular disease. High erythropoietin levels have been demonstrated in the acute erythroblastopenia that occasionally occurs with sickle cell anemia, indicating that the erythropoietic crisis in this condition does not arise from a failure of erythropoietin production.[52]

Polycythemia

A more meaningful classification of the polycythemias based on our knowledge of erythropoietin physiology is now possible. Some cases are secondary to an appropriate increase in erythropoietin production because of hypoxia.[23] The hypoxia may be a result of the atmospheric environment, pulmonary disease, chronic heart disease, vascular disease, or a reduced oxygen carrying capacity of the blood due to methemoglobinemia. Furthermore, several abnormal hemoglobins have been discovered which bind oxygen tightly and do not release it readily to the body tissues, thereby creating a relative tissue hypoxia.[60] In these conditions, when a definite increase in the blood or urinary erythropoietin levels cannot be demonstrated, such patients often respond to phlebotomy with an exaggerated increase in their daily excretion of urinary erythropoietin.[1] This anemia tolerance test stresses the erythropoietin-secreting mechanism just as a glucose tolerance test stresses the pancreas. In the former case the exaggerated increase in urinary erythropoietin excretion probably comes from an organ that has become hyper-

plastic due to previous hypersecretion, while in the latter case the reduced secretion arises from an atrophied endocrine tissue. In both cases, however, the reaction to a particular stress appears to produce a more sensitive test situation than the analysis of a single blood or urinary specimen. A case of familial polycythemia has also been reported in which an exaggerated increase in urinary erythropoietin occurred after phlebotomy.[3] The ultimate cause of the increased secretion in this particular case, however, is not apparent.

Some cases of polycythemia are secondary to an inappropriate increase in erythropoietin production, since no physiological stimulus appears to be present. This increased secretion appears to be autonomous and is associated with the presence of renal disease – hydronephrosis, cysts, adenoma, carcinoma, and renal artery stenosis – cerebellar hemangioblastoma, hepatoma, uterine leiomyoma, and a variety of hyperplasias and tumors of the endocrine organs.[36] In such cases phlebotomy results in no further increase in urinary erythropoietin levels.[1] Whether the polycythemia results from an appropriate or inappropriate increase in erythropoietin secretion, only the number of red cells is increased, since the white cell and platelet counts are unaffected by erythropoietin.

In contrast to these polycythemias is polycythemia vera or primary polycythemia, in which increased erythropoiesis proceeds autonomously without apparent stimulation by erythropoietin. Plasma erythropoietin levels are not elevated in this condition[51] and urinary erythropoietin levels are lower than normal.[1] In addition, the white cell and platelet counts are often increased, since the disease is not limited to erythropoiesis. When patients with polycythemia vera are phlebotomized, only a modest increase in urinary erythropoietin content is observed.[1] This reduced response, compared to the response in patients with secondary polycythemia, may be imposed by prolonged periods of hyposecretion of the hormone due to the abundant red cell mass.

When the marrows of patients with polycythemia vera were tested with erythropoietin in vitro their response was very small in comparison to normal marrows.[35] After a remission of the disease was produced with busulfan, the marrows responded normally to the hormone. This work indicated that two cell lines might be present in polycythemia vera: an abnormal cell line which proliferates autonomously to create a plethora and inhibit erythropoietin production, and a normal cell line which expresses itself after treatment of the abnormal cells and is regulated by the secretion of erythropoietin.[35] The prolonged periods of remission in which a normal hemoglobin is maintained after treatment with ^{32}P or busulfan are thereby accounted for by the re-emergence of the normal cell line which is responsive to the hormone. A disease in mice initiated by the injection of Friend virus filtrates closely resembles polycythemia vera and may prove to be an important model for studying the disease in human beings.[45]

SUMMARY

Erythropoietin has long been thought of as a hormone in search of a disease. As study of the hormone has progressed, however, the possible

usefulness of erythropoietin in certain clinical situations has become more apparent. The anemia which accompanies uremia is the most striking illustration of this possibility. A pure erythropoietin is still not available for clinical trial but should eventually be of use in anemias associated with a low level of the hormone. The work with androgens suggests that pharmacological doses of erythropoietin may be helpful to patients suffering from a variety of anemias which already have increased levels of the hormone. A partially purified erythropoietin preparation is now available for study in animals[53] and this should greatly accelerate investigation of the hormone, including the means of purifying it for human studies, and its possible use in clinical situations. In the meantime investigation of the action of the hormone in laboratory animals and on human marrow cells in vitro should continue to increase our knowledge of the pathogenesis of anemia and polycythemia.

REFERENCES

1. Adamson, J. W.: The erythropoietin/hematocrit relationship in normal and polycythemic man: Implications of marrow regulation. Blood, *32*:597, 1968.
2. Alexanian, R.: Urinary excretion of erythropoietin in normal men and women. Blood, *28*:344, 1966.
3. Alperin, J. B., Levin, W. C., Alexanian, R., and Houston, E. W.: Familial erythrocytosis due to increased erythropoietin production. Clin. Res., *16*:40, 1968.
4. Cole, R. J., and Paul, J.: The effects of erythropoietin on haem synthesis in mouse yolk sac and cultured foetal liver cells. J. Embryol. Exper. Morphol., *15*:245, 1966.
5. Cotes, P. M., and Bangham, D. R.: Bio-assay of erythropoietin in mice made polycythemic by exposure to air at a reduced pressure. Nature, *191*:1065, 1961.
6. DeGowin, R. L., Hofstra, D., and Gurney, C. W.: A comparison of erythropoietin bioassays. Proc. Soc. Exper. Biol. Med., *110*:48, 1962.
7. Erslev, A. J.: Erythropoietin function in uremic rabbits. Arch. Int. Med., *101*:407, 1958.
8. Erslev, A. J.: The effect of anemic anoxia on the cellular development of nucleated red cells. Blood, *14*:386, 1959.
9. Erslev, A. J., McKenna, P. J., Capelli, J. P., Hamburger, R. J., Cohn, H. E., and Clark, J. E.: Rate of red cell production in two nephrectomized patients. Arch. Int. Med., *122*:230, 1968.
10. Finne, P. H.: Erythropoietin in concentrates of urine from healthy persons. Brit. Med. J., *1*:697, 1965.
11. Finne, P. H.: Antenatal diagnosis of the anemia in erythroblastosis. A comparison between spectrophotometry and erythropoietin determination in the amniotic fluid. Acta Paediat. Scandinav., *55*:609, 1966.
12. Finne, P. H.: On placental transfer of erythropoietin. Acta Paediat. Scandinav., *56*:233, 1967.
13. Finne, P. H., and Halvorsen, S.: Transfer of erythropoietin to amniotic fluid in Rh-immunized pregnant women. Sangre, *9*:109, 1964.
14. Fisher, J. W., Roh, B. L., and Halvorsen, S.: Inhibition of erythropoietic effects of hormones by erythropoietin antisera in mildly plethoric mice. Proc. Soc. Exper. Biol. Med., *126*:97, 1967.
15. Fried, W., and Gurney, C. W.: Erythropoietic effect of plasma from mice receiving testosterone. Nature, *206*:1160, 1965.
16. Fried, W., Kilbridge, T., Krantz, S., McDonald, T. P., and Lange, R. D.: Studies on extrarenal erythropoietin. J. Lab. Clin. Med., *73*:244, 1969.
17. Fried, W., Plzak, L., Jacobson, L. O., and Goldwasser, E.: Studies on erythropoiesis. III. Factors controlling erythropoietin production. Proc. Soc. Exper. Biol. Med., *94*:237, 1957.
18. Goldwasser, E.: Biochemical control of erythroid cell development. *In* Monroy, A., and Moscona, A. A., eds.: Current Topics in Developmental Biology. New York, Academic Press, 1966, p. 173.
19. Goldwasser, E., and Kung, C. K. H.: Progress in the purification of erythropoietin. Ann. N.Y. Acad. Sci., *149*:49, 1968.
20. Gordon, A. S., Cooper, G. W., and Zanjani, E. D.: The kidney and erythropoiesis. Seminars in Hematol., *4*:337, 1967.

21. Gross, M., and Goldwasser, E.: Effect of erythropoietin on RNA synthesis. Federation Proc., *27*:394, 1968.
22. Gurney, C. W.: Relationship of erythropoietin production to renal abnormalities. *In* Jacobson, L. O., and Doyle, M., eds.: Erythropoiesis. New York, Grune & Stratton, 1962, p. 359.
23. Gurney, C. W.: Erythropoietin and erythropoiesis. Ann. Int. Med., *65*:377, 1966.
24. Gurney, C. W., and Fried, W.: The regulation of numbers of primitive hemopoietic cells. Proc. Nat. Acad. Sci., *54*:1148, 1965.
25. Gutnisky, A., and Van Dyke, D.: Normal response to erythropoietin or hypoxia in rats made anemic with turpentine abscess. Proc. Soc. Exper. Biol. Med., *112*:75, 1963.
26. Halvorsen, S., and Finne, P. H.: Erythropoietin production in the human fetus and newborn. Ann. N.Y. Acad. Sci., *149*:576, 1968.
27. Hansen, P.: Studies on a method of concentrating erythropoietin from plasma and kidney extract. Scandinav. J. Clin. Lab. Invest., *15*:622, 1963.
28. Hrinda, M. E., and Goldwasser, E.: Facilitated iron entry into marrow cells: An induced permease-like activity. Fed. Proc., *25*:284, 1966.
29. Jacobson, L. O., Goldwasser, E., Fried, W., and Plzak, L.: Role of the kidney in erythropoiesis. Nature, *179*:633, 1957.
30. Jacobson, L. O., Goldwasser, E., Gurney, C. W., Fried, W., and Plzak, L.: Studies of erythropoietin: The hormone regulating red cell production. Ann. N.Y. Acad. Sci., *77*:551, 1959.
31. Jacobson, L. O., Goldwasser, E., Plzak, L., and Fried, W.: Studies on erythropoiesis. IV. Reticulocyte response of hypophysectomized and polycythemic rodents to erythropoietin. Proc. Soc. Exper. Biol. Med., *94*:243, 1957.
32. Jacobson, L. O., Marks, E. K., and Gaston, E. O.: Studies on erythropoiesis. XII. The effect of transfusion-induced polycythemia in the mother on the fetus. Blood, *14*:644, 1959.
33. Jepson, J. H., and Lowenstein, L.: Panhypoplasia of the bone marrow. I. Demonstration of a plasma factor with anti-erythropoietin-like activity. Canad. Med. Assoc. J., *99*:99, 1968.
34. Keighley, G.: Further experiences with assays, units and standards of erythropoietin. Ann. N.Y. Acad. Sci., *149*:18, 1968.
35. Krantz, S. B.: Response of polycythemia vera marrow to erythropoietin *in vitro*. J. Lab. Clin. Med., *71*:999, 1968.
36. Krantz, S. B., and Jacobson, L. O.: Erythropoietin and the Regulation of Erythropoiesis. Chicago, University of Chicago Press, 1970.
37. Krantz, S. B., Gallien-Lartigue, O., and Goldwasser, E.: The effect of erythropoietin upon heme synthesis by marrow cells *in vitro*. J. Biol. Chem., *238*:4085, 1963.
38. Krantz, S. B., and Goldwasser, E.: On the mechanism of erythropoietin-induced differentiation. II. The effect on RNA synthesis. Biochim. Biophys. Acta, *103*:325, 1965.
39. Krantz, S. B., and Kao, V.: Studies on red cell aplasia. I. Demonstration of a plasma inhibitor to heme synthesis and an antibody to erythroblast nuclei. Proc. Nat. Acad. Sci., *58*:493, 1967.
40. Krzymowski, T., and Czarnypzewicz, J.: Plasma erythropoietin in physiological conditions. Acta Physiol. Polon., *15*:140, 1964.
41. Kuratowska, Z.: On the subcellular localization of the erythropoietic renal factor. Bull. Acad. Polon. Sci., *13*:385, 1965.
42. Lewis, J. P., and Trobaugh, F. E., Jr.: Haematopoietic stem cells. Nature, *204*:589, 1964.
43. Lowy, P. H., and Keighley, G.: Extraction of erythropoietin from human urine and rabbit plasma. Clin. Chim. Acta, *13*:491, 1966.
44. McKenzie, D., Friedman, R., Katz, S., and Lanzkowsky, P.: Erythropoietin levels in anaemia in kwashiorkor. S. African Med. J., *41*:1044, 1967.
45. Mirand, E. A.: Virus-induced erythropoiesis in hypertransfused-polycythemic mice. Science, *156*:832, 1967.
46. Mirand, E. A., Gordon, A. S., and Wenig, J.: Mechanism of testosterone action in erythropoiesis. Nature, *206*:270, 1965.
47. Mirand, E. A., Murphy, G. P., Steeves, R. A., Weber, H. W., and Retief, F. P.: Extra-renal production of erythropoietin in man. Acta Haemat., *39*:359, 1968.
48. Moores, R. R., Wright, C.-S., Collins, L. R., Gardner, E., Jr., Lewis, J. P., and Smith, L. L.: Stimulation of erythropoietin by androgens in the human. Scandinav. J. Haemat., *5*:415, 1968.
49. Naets, J. P.: Erythropoiesis in nephrectomized dogs. Nature, *181*:1134, 1958.
50. Naets, J. P., and Wittek, M.: Presence of erythropoietin in the plasma of one anephric patient. Blood, *31*:249, 1968.
51. Noyes, W. D., Domm, B. M., and Willis, L. C.: Regulation of erythropoiesis. I. Erythropoietin assay as a clinical tool. Blood, *20*:9, 1962.
52. Okcuoglu, A.: Erythropoietin in acute erythroblastopenia of sickle cell anemia. Turk. J. Pediat., *6*:225, 1964.

53. Painter, R. H., Bruce, W. R., and Goldwasser, E.: The commercial production of erythropoietin from anemic sheep plasma. Ann. N.Y. Acad. Sci., *149*:71, 1968.
54. Paul, J., and Hunter, J. A.: DNA synthesis is essential for increased haemoglobin synthesis in response to erythropoietin. Nature, *219*:1362, 1968.
55. Rishpon-Myerstein, N., Kilbridge, T., Simone, J., and Fried, W.: The effect of testosterone on erythropoietin levels in anemic patients. Blood, *31*:453, 1968.
56. Schooley, J. C.: Responsiveness of hematopoietic tissue to erythropoietin in relation to the time of administration and duration of action of the hormone. Blood, 25:795, 1965.
57. Shalet, M., Coe, D., and Reissmann, K. R.: Mechanism of erythropoietic action of thyroid hormone. Proc. Soc. Exper. Biol. Med., *123*:443, 1966.
58. Stohlman, F., Jr., Beland, A., and Howard, D.: Mechanism of macrocytic response to erythropoietin. J. Clin. Invest., *42*:984, 1963.
59. Till, J. E., and McCulloch, E. A.: A direct measurement of the radiation sensitivity of normal mouse bone marrow cells. Radiation Res., *14*:213, 1961.
60. Weatherall, D. J.: Polycythemia resulting from abnormal hemoglobins. New Eng. J. Med., *280*:604, 1969.

Clinical Pathology Department
National Institutes of Health
Bethesda, Maryland 20014

Glucagon in Medicine: New Ideas for an Old Hormone

A. M. LAWRENCE, M.D., Ph.D.*

Glucagon was discovered serendipitously in 1923 as a contaminant of early insulin preparations.[6] Since that time, a welter of observations and facts have been accumulated on which current day assessment of glucagon's place in clinical medicine must be made. Unlike its companion hormone, insulin, glucagon's physiologic role has been an elusive one to document, particularly in man. In large measure this has been due to inability to recognize clinical states of excess or deficiency, to inability to produce an experimental model of glucagon deficiency to equal or even rival the alloxan-diabetic counterpart, and finally to insensitive assay techniques. This latter difficulty has been obviated by the recent advent of sensitive radioimmunoassays, which have made it possible to measure picagram quantities of immunoassayable glucagon in biologic fluids, thereby offering an invaluable tool for furthering our understanding of glucagon's role in mammalian physiology.

THE PHYSIOLOGIC ROLE OF GLUCAGON

Substrate Provider

Glucagon is a peptide hormone consisting of 29 amino acids with a molecular weight of 3400. It originates in the alpha cells of the islets of Langerhans. It is best known to traditional students of physiology as the hyperglycemic-glycogenolytic principle extracted from the pancreas or from isolated islets of Langerhans.[9] The hyperglycemic effect of glucagon is due to increased hepatic glucose output secondary to stimulation of 3′5′-AMP formation, activation of hepatic phosphorylase, and ensuing hepatic glycogenolysis.[26, 27]

Although earlier studies of the physiologic effects of glucagon emphasized its glycogenolytic action and suggested its particular need in real or imminent glycopenic situations, subsequent study has indicated a potentially broader role of this hormone in securing a variety

Associate Professor of Medicine (Endocrinology), Department of Medicine, University of Chicago, Chicago, Illinois

of readily available energy substrates and patterns thereof. Glucagon is gluconeogenetic. It lowers plasma alpha amino acids and stimulates hepatic conversion of amino acids to glucose and glycogen[10] and may promote utilization of short chain free fatty acids for glucose production by the liver.[34] Glucagon is clearly lipolytic in certain species such as fowl[11] and probably promotes lipolysis in man as well. The in vivo lipolytic effects of glucagon in man are largely masked by the lipogenic effects of insulin secreted in response to the hyperglycemic and insulino-secretory effects of glucagon.

The Entero-Insular Axis

With the reports of Roger Unger and his associates in 1959–1961,[29] the radioimmunoassay era for glucagon was ushered in. Since that time a body of data has been collected, primarily from dog experiments, which lends strong support to an earlier concept that important metabolic interplay exists between the islets of Langerhans and humoral factors originating from the digestive tract and its exocrine appendages.

As far back as 1947, it was reported that blood obtained from normal man following oral glucose administration would cause hyperglycemia in human volunteers, and it was believed that the hyperglycemia-promoting substance was probably glucagon.[5] In 1948, Sutherland and deDuve demonstrated a material in extracts of stomach and intestinal mucosa which promoted hepatic glucose output.[25] Subsequent demonstration that these extracts also promoted accumulation of hepatic cyclic AMP strongly indicated that the substance was indeed "glucagon-like."[15] Later it was demonstrated that this material was immunologically similar to pancreatic glucagon. It is currently believed that there may be more than one glucagon-like material extractable from gut mucosa immunologically reactive with antiglucagon serum, possessing some but not all of pancreatic glucagon's biological activities.[31]

These observations gained added meaning when it could be shown by radioimmunoassay that glucagon can promote insulin release,[19] that large oral glucose[14, 20] and amino acid loads[2] raise circulating titers of immunoassayable glucagon, and that insulin output in response to oral glucose is much greater than to equivalent hyperglycemia by the intravenous route.[7] Finally it was shown that serum glucagon elevations in response to oral glucose are of digestive tract origin,[31] whereas immunoassayable glucagon rises to intrajejunal instillation or intravenous administration of amino acids are of pancreatic origin.[16] Thus it seems likely that there are at least two glucagons, one pancreatic, the other of enteric origin.

The seeming paradox of glucagon as an insulino-secretory signal and a promoter of hepatic glucose output is resolved by considering that glucagon release, enteric or pancreatic in origin, allows for a more propitious early release of insulin at the time of meal consumption. In this manner insulin release need not wait for alimentary hyperglycemia, a time lag which might lead to a greater quantity of absorbed glucose reaching the periphery than would be warranted by prevailing peripheral energy requirements. Sufficient insulin release, in turn, acts to counter

the acute glycogenolytic effects of glucagon at a time when hepatic storage of glucose may be indicated.

This intriguing scheme is far more complex, however, and appears to involve other gastrointestinal hormonal mediators of early insulin release. Gastrin, secretin, and pancreozymin, as well as enteroglucagon, stimulate insulin release, and pancreozymin has been shown, in dogs, to stimulate pancreatic glucagon release as well. Finally, some studies have reported the presence of substantial quantities of insulin in bile, while other data support the proposition that insulin present in the duodenum acts to enhance glucose absorption.[12, 28] Thus, given that these observations, made in dogs, can be extended to man, a complex spectrum of interplay exists between the alimentary system and the islets of Langerhans which serves to effectively modulate the disposition of absorbed foodstuffs in a manner most appropriate to the particular and shifting substrate needs of specialized organs and tissues of the body.

Regulation of Secretion

The radioimmunoassay era has also yielded more precise information on factors regulating the release or secretion of glucagon. This information has been derived from the study of peripheral venous sampling in man, pancreatic and jejunal vein sampling in dogs, and sampling of incubation media bathing isolated islets of Langerhans.

Peripheral venous glucagon titers in man rise following ingestion of glucose,[14, 20] protein, or amino acids,[2] and particularly after the intravenous administration of arginine or other mixtures of essential amino acids.[2] Glucagon levels may rise during starvation, but this point is still controversial. Acute hypoglycemia in animals is accompanied by rises in pancreatic vein glucagon.[4] Hyperglycemia suppresses high glucagon titers,[30] and modest suppression has been noted in pancreatic vein blood during heparin-induced rises in plasma free fatty acids.[22] Since the recognition of immunologically reactive digestive-tract glucagon, studies in dogs and with isolated islets have shown that the rise following oral glucose is of enteric origin,[30] while that following protein and amino acids is of pancreatic origin.[16] Although all the enteric hormones studied to date stimulate insulin release, only pancreozymin has caused pancreatic glucagon release in dogs.[28]

THE CLINICAL IMPORTANCE OF GLUCAGON

Glucagon Deficiency

Suspected cases of glucagon deficiency have hitherto been based on conjecture, i.e., fasting hypoglycemia coupled with light microscopic evaluation of islet morphology. Basal levels of peripheral venous immunoassayable glucagon in man are low or below the sensitivity of available assay systems in many normal individuals. Recently, using a particularly sensitive and discriminatory antiserum, not reactive with enteroglucagon, Unger has proposed that arginine infusion be used as a provocative test of glucagon reserves; failure to induce a measured rise in immunoassayable glucagon would serve to furnish evidence of alpha

cell failure.[1] Thus it may be possible, via this approach, to establish whether certain states of idiopathic hypoglycemia, reactive hypoglycemia, chronic relapsing pancreatitis with fasting hypoglycemia, brittle diabetes mellitus, and the dumping syndrome are associated with glucagon deficiency or not. Unger and his group have presented preliminary data to show that selected patients with chronic pancreatitis and fasting hypoglycemia fail to evince a normal glucagon rise with intravenous arginine administration.[1]

Hyperglucagonemia

Immunoassayable elevation in circulating titers of glucagon in man has been demonstrated in a patient with a functioning alpha cell carcinoma, in some pancreatic tumor extracts, in a hepatic metastasis of a bronchogenic carcinoma, in some patients with acute pancreatitis, in a few cases of pheochromocytoma, in many patients with hyperparathyroidism, in patients and kindred of patients with the multiple endocrine adenomatosis syndrome, and in some severely obese, mildly diabetic children with hyperinsulinemia.[12] There is as yet no support for the speculation that hyperglucagonemia may be a causative factor in some forms of diabetes mellitus, however.

The Chromaffin System and the Heart

Important effects of glucagon upon the sympathetic nervous system and upon the heart have come to light since the report of Scian and co-workers in 1960.[21] These workers demonstrated that catecholamine release by isolated dog adrenal glands can be strikingly increased by glucagon perfusion. Confirmation of this data in other animals has been obtained, and in man glucagon has proved to be a valuable addition to the armamentarium of provocative drugs for selected cases of suspected pheochromocytoma.[13] Striking release of stored catechols from chromaffin tumors can be obtained with 0.5 to 1 mg. of glucagon administered intravenously; the potential advantages of glucagon over histamine are the absence of histamine-induced hypotension, headache, increased cerebrospinal fluid pressure, and bronchospasm.

In normal man, intravenous glucagon exerts a positive chronotropic effect upon the sinus node. Glucagon-induced tachycardia can be partially but not wholly blocked by reserpine or beta-adrenergic blockade, and it is therefore believed to act directly upon the sinus node as well as probably stimulating local sympathetic discharge of stored norepinephrine.[33]

In addition to its chronotropic effects and its action upon catecholamine discharge, glucagon appears to exert an important influence upon electrical activity of the heart as well as upon myocardial function. The net effect of glucagon on myocardial performance is one of positive inotropism. Cardiac output and rate of rise of left ventricular pressure increase. This occurs even in the fully digitalized and reserpinized animal and under beta-adrenergic blockade.[8] Glucagon also decreases atrioventricular conduction time in dogs and man and may increase myocardial irritability.

Because of these positive effects, glucagon has been proposed as a potential therapeutic agent in patients with usual etiologies for failing myocardial performance and for cardiac depression secondary to beta-adrenergic blockade with propanolol.[18] Initial enthusiasm, however, has waned a bit, as not all laboratories have been able to confirm the initial promising results obtained in man. Although glucagon does promote natriuresis, this diuresis is short lived. Increased myocardial irritability and arrhythmias have been reported and severe nausea attending intravenous infusion in some patients may vitiate its full therapeutic usefulness. The data, nevertheless, are provocative and are providing experimental and clinical cardiologists with an exciting experimental and potentially useful cardiotonic drug. Whether glucagon or a congener of this peptide will find acceptance in therapeutic cardiology remains to be proved, however.

Glucagon and the Digestive Tract

The embryologic offshoot of the digestive tract out of which the exocrine pancreas develops and from whose ductular elements the islets of Langerhans take origin has long suggested intriguing interplay between these organs in health and disease. The effects of intramuscular or intravenous glucagon upon the digestive tract are interference with gastric, jejunal, and colonic motility, inhibition of gastric secretion, and decrease of pancreatic enzyme secretion.[12] These data suggest that if glucagon is released in response to eating it can promote a feeling of satiety. Further, there is a potential servomechanism called into play whereby meat or amino acids initiate pancreozymin release which then stimulates glucagon release. Glucagon release may then halt pancreatic exocrine secretion at a time when it is no longer needed, thereby providing a neatly programmed mechanism for dealing with on-off-on needs of certain digestive enzymes.

Glucagon and Other Endocrines

Exogenously administered glucagon stimulates insulin secretion[19] and release of adrenal medullary catecholamines.[21] The latter act to blunt insulin release, and this in turn accounts for the mild diabetes mellitus seen in pheochromocytoma patients. Whether this apparent modulated interplay of glucoregulatory hormones normally obtains is not known.

Some relation between glucagon and parathyroid function seems apparent from studies in man.[12] Many patients with hyperparathyroidism have hyperglucagonemia. Glucagon administration to rabbits results in parathyroid hyperplasia. The association of hyperparathyroidism, chronic relapsing pancreatitis, alpha cell hyperplasia, and hyperglucagonemia has been reported.[17] Intravenous administration of glucagon to man results in lowering of the serum calcium. Although glucagon does promote modest hypercalciuria, recent reports seriously suggest that its calcium-lowering effect may be directly related to release of the calcium-lowering peptide calcitonin.[3]

The relationship of glucagon secretion to the syndrome of multiple endocrine adenomatosis is receiving increased attention. In addition to

the hyperglucagonemia noted in many cases of hyperparathyroidism and in pheochromocytoma, recent studies have demonstrated hyperglucagonemia in patients with the Zollinger-Ellison syndrome and kindred with this syndrome or with the syndrome of multiple endocrine adenomas.

THERAPEUTIC AND DIAGNOSTIC USES OF GLUCAGON

Hypoglycemia

Glucagon, 1 mg. subcutaneously, can be used effectively to shorten insulin-induced hypoglycemia and is of particular use in those brittle diabetics in whom hypoglycemia attacks are attended by mental confusion or coma so that oral intake of food is difficult. Family members can be taught to administer glucagon, and attacks of unconsciousness secondary to hypoglycemia can be reversed within 20 to 30 minutes after glucagon administration. Although recent reports demonstrate the stimulating effect of glucagon upon insulin release, this action of glucagon is negligible or absent in the insulin-deficient diabetic.

Glycogen Storage Disease

Glucagon and especially the longer acting zinc glucagons have been useful in some forms of glycogen storage disease in which reduction in liver size occurs secondary to activation of phosphorylase and ensuing glycogenolysis. It may serve a twofold usefulness in such patients by raising the fasting blood sugar as well as diminishing liver size.[24]

Other Uses

Other potential therapeutic uses of glucagon are several, but are still undergoing experimental evaluation. Its inhibition of gastric and pancreatic secretion has suggested a potential use in ulcer disease and acute pancreatitis. Although it will lower the serum calcium toward or to normal in certain hypercalcemic patients, large doses must be given by continuous intravenous infusion, limiting its over-all usefulness if the hypercalcemia poses a chronic problem. Glucagon's positive inotropic effect coupled with its natriuretic action have suggested its usefulness in patients with declining myocardial function and congestive heart failure. Reports of increased myocardial irritability and its chronotropic action lessen the likelihood that glucagon will pose as a perfect cardiotonic drug.

Diagnostic

Glucagon is used diagnostically as a provocative agent in selected patients with suspected pheochromocytoma when such tests seem warranted.[13] It is useful in certain forms of glycogen storage disease, where failure to cause hyperglycemia in a fed patient connotes defective hepatic glycogenolysis.[24] In insulinoma patients, 1 mg. intravenously may cause hypoglycemia rather than the expected rise in blood sugar. Glucagon may prove useful as a diagnostic agent when deficient adrenal

medullary function is suspected, and failure to obtain a rise in circulating catecholamines following intravenous glucagon implies a nonresponsive adrenal medulla.

Side-Effects

Glucagon's minor but potentially hazardous side effects must be kept in mind when considering glucagon for either diagnostic or therapeutic effects. Need for injection, intestinal atony, nausea and vomiting, catecholamine discharge and myocardial irritability, insulin release and hypoglycemia, and protein allergic reactions require recognition in this regard.

REFERENCES

1. Aguilar-Parada, E., Eisentraut, A. M., and Unger, R. H.: Pancreatic glucagon secretion in normal and diabetic subjects. Amer. J. Med. Sci., 257:415-419, 1969.
2. Assan, R., Rosselin, G., and Dolias, J.: Effects sûr la glucagonemiedes perfusions et ingestions d'acides amines. Jornées Annuelles de Diabetologie de l'Hôtel-Dieu, 7:25-41, 1967.
3. Avioli, L. V., Birge, S. J., Scott, S., and Shieber, W.: Role of the thyroid gland during glucagon-induced hypocalcemia in the dog. Amer. J. Physiol., *216*:939-945, 1969.
4. Buchanan, K. B., Vance, J. E., Dinstl, K., and Williams, R. H.: Effect of blood glucose on glucagon secretion in anesthetized dogs. Diabetes, *18*:11-18, 1969.
5. Burger, M., and Klotzbucher, E.: Der Nachweis zuckerregelnder Hormone (Insulin, Glukagon) im menschlichen Blute durch Transfusion. Zeitschr. Ges. Inn. Med., 2: 43-50, 1947.
6. Collip, J. B.: Delayed manifestation of the physiologic effects on insulin following administration of certain pancreatic extracts. Amer. J. Physiol., *63*:391-392, 1923.
7. Elrick, H., Stimmler, L., Hlad, C. J., and Arai, Y.: Plasma insulin response to oral and intravenous glucose administration. J. Clin. Endocrinol., *24*:1076-1082, 1964.
8. Farah, A., and Tuttle, R.: Studies on the pharmacology of glucagon. J. Pharmacol. Exper. Therap., *129*:49-55, 1960.
9. Foa, P. P.: Glucagon, the hyperglycemic-glycogenolytic hormone of the pancreas. Adv. Int. Med., *6*:29-58, 1954.
10. Garcia, A., Williamson, J. R., and Cahill, G. F.: Studies on the perfused rat liver. II. Effect of glucagon on gluconeogenesis. Diabetes, *15*:188-193, 1966.
11. Grande, F.: Effect of glucagon on plasma free fatty acids and blood sugar in birds. Proc. Soc. Exp. Biol. Med., *128*:532-536, 1968.
12. Lawrence, A. M.: Glucagon. Ann. Rev. Med., *20*:207-223, 1969.
13. Lawrence, A. M.: Glucagon provocative test for pheochromocytoma. Ann. Int. Med., *66*:1091-1096, 1967.
14. Lawrence, A. M.: Radioimmunoassayable glucagon levels in man. Effects of starvation, hypoglycemia, and glucose administration. Proc. Nat. Acad. Sci., *55*:316-320, 1966.
15. Makman, M. H., and Sutherland, E. W.: Use of liver adenyl cyclase for assay of glucagon in human gastro-intestinal tract and pancreas. Endocrinol., *75*:127-134, 1964.
16. Ohneda, A., Parada, E., Eisentraut, A. M., and Unger, R. H.: Characterization of response of circulating glucagon to intraduodenal and intravenous administration of amino acids. J. Clin. Invest., *47*:2305-2322, 1968.
17. Paloyan, E., Lawrence, A. M., Straus, F. H., Paloyan, D., Harper, P. V., and Cummings, D. V.: Alpha cell hyperplasia in calcific pancreatitis associated with hyperparathyroidism. J.A.M.A., *200*:757-761, 1967.
18. Parmley, W. W., Glick, G., and Sonnenblick, E. H.: Cardiovascular effects of glucagon in man. New Eng. J. Med., 279:12-17, 1968.
19. Samols, E., Marri, G., and Marks, V.: Promotion of insulin secretion by glucagon. Lancet, 2:415-416, 1965.
20. Samols, E., Marri, G., Tyler, J., and Marks, V.: Stimulation of glucagon secretion by oral glucose. Lancet, 2:1257-1259, 1965.
21. Scian, L. F., Westermann, C. D., Verdesca, A. S., and Hilton, J. G.: Adrenocortical medullary effects of glucagon. Amer. J. Physiol., *199*:867-870, 1960.
22. Seyffert, W. A., Jr., and Madison, L. L.: Physiologic effects of metabolic fuels on carbohydrate metabolism. Diabetes, *16*:765-766, 1967.

23. Sokol, J. E.: Effect of glucagon on gluconeogenesis by the isolated perfused rat liver. Endocrinol., *78*:538-548, 1966.
24. Sokol, J. E., Lowe, C. U., and Sarcione, E. J.: Liver glycogen disease. Arch. Int. Med., *109*:612-624, 1962.
25. Sutherland, E. W., and deDuve, C.: Origin and distribution of hyperglycemia-glycogenetic factor of pancreas. J. Biol. Chem., *175*:663-674, 1948.
26. Sutherland, E. W., and Rall, T. W.: The relation of adenosine-3′,5′-phosphate and phosphorylase to the action of catechol amines and other hormones. Pharmacol. Rev., *12*:265-299, 1960.
27. Sutherland, E. W., Robison, G. A., and Butcher, R. W.: Some aspects of the biological role of adenosine 3′,5′-monophosphate (cyclic AMP). Circulation, *37*:279-306, 1968.
28. Unger, R. H.: New ideas concerning the physiologic role of glucagon. Amer. J. Med. Sci., *255*:273-276, 1968.
29. Unger, R. H., Eisentraut, A. M., McCall, M. S., and Madison, L. L.: Glucagon antibodies and an immunoassay for glucagon. J. Clin. Invest., *40*:1280-1289, 1961.
30. Unger, R. H., Ohneda, A., Valverde, I., Eisentraut, A. M., and Exton, J.: Characterization of the responses of circulating glucagon-like immunoreactivity to intraduodenal and intravenous administration of glucose. J. Clin. Invest., *47*:48-65, 1968.
31. Valverde, I., Rigopoulou, D., Exton, J., Ohneda, A., Eisentraut, A., and Unger, R. H.: Demonstration and characterization of a second fraction of glucagon-like immunoreactivity in jejunal extracts. Amer. J. Med. Sci., *255*:415-420, 1968.
32. Vance, J. E., Stoll, R. W., Kitabchi, A. E., Williams, R. H., and Wood, F. C., Jr.: Nesidioblastosis in familial endocrine adenomatosis. J.A.M.A., *207*:1679–1682, 1969.
33. Whitehouse, F. W., and James, T. N.: Chronotropic action of glucagon on the sinus node. Proc. Soc. Exp. Biol. Med., *122*:823-826, 1960.
34. Williamson, J. R.: Mechanism for the stimulation in vivo of hepatic gluconeogenesis by glucagon. Biochem. J., *101*:11C-14C, 1966.

950 East 59th Street
Chicago, Illinois 60637

Proinsulin

The Single Chain Precursor of Insulin

ARTHUR H. RUBENSTEIN, M.D., F.C.P.(S.A.), M.R.C.P.(Lond.)*

DONALD F. STEINER, M.D.**

There is a growing body of evidence to support the concept that diabetes mellitus is characterized by impaired insulin secretion from the pancreatic beta cells. The elegant studies of Kipnis,[11] Cerasi and Luft,[2] and Colwell and Lein[8] have indicated that plasma immunoreactive insulin levels in prediabetics and diabetics are lower than in carefully matched control subjects. Furthermore, the pathological findings of beta cell destruction and decreased quantities of extractable insulin in the pancreas of juvenile diabetics are consistent with this view.[30] It is thus not surprising that a great deal of effort has been expended in the study of the biosynthesis, storage, and release of insulin under the influence of a variety of stimuli. In this regard the recent discovery of proinsulin has provided new opportunities for understanding the functional activity and regulatory mechanism of beta cells.

PROINSULIN AND THE BIOSYNTHESIS OF INSULIN

The existence of proinsulin, the single polypeptide chain precursor of insulin, was first demonstrated in studies of insulin biosynthesis in human islet cell adenomas.[32, 34] When slices of these tumors were incubated with labeled amino acids, radioactivity was first incorporated into a protein of higher molecular weight, which was, however, immunologically related to insulin. It consisted of the A and B chains of insulin linked by an additional polypeptide segment. It could be transformed to an insulin-like component by tryptic hydrolysis. Subsequent experiments with islet tissue from the rat,[7, 32] codfish,[12] and fetal calf[35] have confirmed the presence of such a proinsulin in the biosynthetic pathway

*Assistant Professor, Department of Medicine, University of Chicago, Chicago, Illinois
**Professor, Department of Biochemistry, University of Chicago, Chicago, Illinois

leading to insulin. Studies with isolated rat islets have indicated that newly synthesized insulin was first detected after a delay of 15 to 25 minutes and that its rate of production then became linear, while proinsulin production tended to plateau. The half-life for the transformation of proinsulin to insulin was approximately 60 minutes.[29] Newly synthesized insulin was secreted from the islets after 1 hour's incubation and small amounts of proinsulin appeared in the medium along with the labeled insulin.[29]

The question as to the site of transformation of proinsulin to insulin is an important one. Jamieson and Palade[13] have shown that secretory proteins in the exocrine pancreatic cells are synthesized in the rough endoplasmic reticulum. They are then sequestered in the cisternae of the endoplasmic reticulum which communicate with the Golgi apparatus. Packaging of the protein into secretory granules takes place in proximity to this organelle. The whole process requires approximately an hour. A similar progression of labeled protein in beta cells has now been found by Lacy,[15] and this evidence combined with our kinetic data suggests that the transformation process is intimately associated with granule formation by vesiculation from the Golgi periphery. The enzyme(s) responsible for the conversion of proinsulin to insulin are probably added or activated in the Golgi region, but as yet have not been isolated.

The precise role of the connecting segment in the proinsulin molecule is still open to speculation. However, it appears that its major function is to facilitate the efficient formation of the disulfide linkages between the two insulin chains and to ensure their correct alignment during biosynthesis. This hypothesis has been strengthened by the demonstration that after reduction in 8 M urea, reoxidation of proinsulin results in the reappearance of 60 to 70 per cent of its native immunologic reactivity, whereas insulin under similar conditions does not refold to greater than 1 or 2 per cent.[1, 4, 30]

When rat islets were incubated with labeled amino acids which appear in the connecting segment of proinsulin, radioactivity was also detected in a free connecting peptide fragment, which we have designated the C-peptide. This peptide, lacking only the terminal basic residues which join it to the A and B chains in the intact proinsulin molecule, was retained within the beta cell along with insulin in essentially equivalent amounts after being cleaved from proinsulin.[7] Moreover, the labeled C-peptide was secreted with newly synthesized insulin into the medium, and the fraction remaining in the islets declined only slightly relative to insulin over a period of 16 hours.[23] A similar peptide[6] has recently been isolated from extracts of bovine and human pancreas and the former has been shown to have an amino acid composition identical to that of the bovine proinsulin connecting segment, but lacking the terminal basic residues. It thus seems probable that the C-peptide is largely retained in the secretory granules after transformation of proinsulin and is then liberated along with insulin into the circulation.[23]

ISOLATION AND CHARACTERIZATION OF PROINSULIN

As proinsulin and insulin tend to crystallize together, commercial insulin preparations have been used a a source of this material.[5, 27, 31, 33]

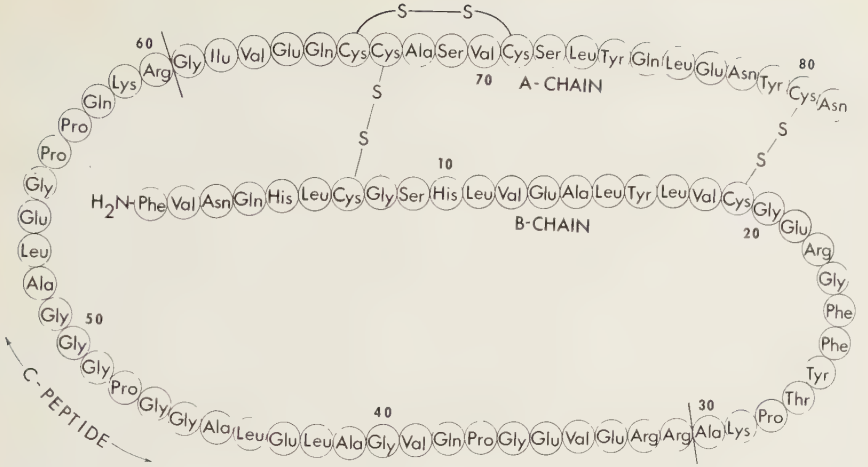

S—S
S
S
S
S
S
60
70
80
10
20
30
40
50
A-CHAIN
B-CHAIN
H_2N
C-PEPTIDE
Gly Ile Val Glu Gln Cys Cys Ala Ser Val Cys Ser Leu Tyr Gln Leu Glu Asn Tyr Cys Asn
Phe Val Asn Gln His Leu Cys Gly Ser His Leu Val Glu Ala Leu Tyr Leu Val Cys Gly Glu Arg Gly Phe Phe Tyr Thr Pro Lys Ala Arg Arg Glu Val Glu Val Gln Pro Gly Glu Val Gly Ala Leu Glu Leu Ala Gly Gly Gly Pro Gly Gly Gly Ala Gly Leu Glu Leu Ala Leu Glu Gln Lys Arg

Figure 1. Structure of bovine proinsulin.

We have separated a crude proinsulin preparation containing three major components from first crystals of bovine insulin by gel filtration. The slowest migrating component on polyacrilamide gel electrophoresis is an intact single chain protein corresponding to the proinsulin isolated in biosynthetic studies, while the second fraction is a mixture of two-chain intermediate forms. Both proinsulin and the intermediate forms are converted to dealanated insulin after short incubation with trypsin. The third component appears to be a dimer of insulin which does not undergo trypsin transformation to insulin (nonconvertible fraction).[33]

The proinsulin molecule consists of a single polypeptide chain of molecular weight about 9000, beginning at the N-terminus with the B-chain sequence of insulin, terminating with the A-chain sequence and bearing in between an additional polypeptide of 30 (bovine)[18] or 33 (porcine)[5] amino acids connecting the A and B chains (Fig. 1). At either end of the connecting segment of proinsulin (C-peptide) are two basic residues which represent the sites of cleavage during the formation of insulin. In contrast to the close similarity of porcine and bovine insulin, the amino acid sequences of the proinsulins differ by 30 per cent. The tentative amino acid composition of the connecting segment of cod proinsulin[12] differs markedly from that in porcine or bovine proinsulin, and the human C-peptide also has several distinctive features.[6] These remarkable species variations suggest that the connecting segment may have few specific structural requirements.

IMMUNOLOGICAL PROPERTIES[25, 31]

Using an insulin antiserum and a tracer of ^{131}I-insulin, proinsulin reacted about half as well as insulin. The bovine intermediate forms and the nonconvertible fraction have greater reactivity than proinsulin. Antisera prepared against purified bovine proinsulin discriminated well between proinsulin and its intermediate forms on the one hand and insulin and the nonconvertible fraction on the other. The presence of anti-

bodies directed specifically against the connecting segment was demonstrated by the inability of large quantities of insulin to compete with ^{131}I-proinsulin for all the binding sites of the antiserum. As the cross reaction with insulin in certain antisera was weak, preincubation of these proinsulin antisera with insulin resulted in a system which selectively measured proinsulin and was not influenced by added insulin. Yip and Logothetopoulos[37] adsorbed their proinsulin antiserum with an immunosorbent of insulin conjugated to Sephadex and succeeded in removing those antibodies which cross reacted with insulin. The resulting antiserum was highly specific for proinsulin. Although its reactivity varied with different antisera, the free C-peptide decreased the ^{131}I-proinsulin bound to proinsulin antiserum. This immunological activity of the C-peptide has important implications, for if it exists free in the beta cell or circulation, it will react with proinsulin antiserum in a manner that will be difficult to distinguish from intact proinsulin.

When bovine, porcine, or human proinsulin was measured using an antiserum to bovine proinsulin and bovine ^{131}I-proinsulin, only bovine proinsulin displaced the tracer effectively.[22] Similarly, only porcine proinsulin gave a steep slope when porcine proinsulin antiserum was used. These results indicate considerable immunological differences between the three species of proinsulin, and specific antisera to each will thus be needed for optimal assay specificity.[24]

BIOLOGICAL ACTIVITY

There is considerable interest in the biological activity of proinsulin and widely divergent results have been reported depending upon the method selected for its assay. Using the isolated fat cell technique,[7] purified bovine proinsulin has about 2 to 5 per cent of the activity of insulin[3, 10] when measured by its effect on glucose-1-^{14}C conversion to $^{14}CO_2$ and the inhibition of glycerol release stimulated by theophylline. Its activity is also very low (less than 5 per cent that of insulin) in the isolated frog sartorius muscle assay.[17] After brief exposure of the proinsulin to trypsin in vitro prior to these assays, its biological activity rises to almost 18 units per mg. However, spontaneous activation of proinsulin in the incubation medium has not been observed in either bioassay system. Challoner has measured the disappearance of immunoassayable insulin and proinsulin from the medium in which isolated fat cells were incubated, and found that whereas the insulin level decreased by 30 to 50 per cent, that of proinsulin was unchanged.[3]

The activity of porcine and bovine proinsulin on rat epididymal fat tissue and diaphragm muscle is approximately 20 to 35 per cent that of insulin when compared on a molar basis.[19, 28] However, the maximal stimulatory effects of proinsulin and insulin are similar. Proinsulin stimulates the same metabolic parameters as insulin in fat tissue, increasing glucose uptake from the medium, its oxidation to carbon dioxide and its incorporation into triglyceride and glycogen. Moreover, proinsulin, like insulin, enhances the oxidation of glucose-1-^{14}C to $^{14}CO_2$ to a

greater extent than it affects oxidation of carbon 6. Both these proteins also stimulate the uptake of glucose and formation of glycogen in the isolated hemidiaphragm.[19, 28] These effects of proinsulin are all completely blocked by the addition of anti-insulin serum and partially inhibited by the Kunitz pancreatic trypsin inhibitor[28] which is known to inhibit the activity of trypsin, chymotrypsin, plasmin, and kallikrein.

On the basis of these observations, Shaw and Chance[28] have suggested that proinsulin has little or no intrinsic biological activity, but that an enzyme or enzyme system present in adipose tissue and muscle converts it to a molecule which has biological and immunological properties similar to insulin. Presumably these enzymes are located in the extracellular compartment of adipose tissue and are removed by collagenase digestion, which is a necessary step in the preparation of isolated fat cells. In this regard, we have injected iodinated proinsulin into rats to determine whether it is converted to an insulin-like component in the circulation. By bleeding the animals at varying times and gel filtering aliquots of the serum with carrier insulin, we have been unable to observe any transformation to an insulin-like component. These results appear to rule out massive intravascular transformation, but do not exclude the possibility that proteolytic activation may occur in peripheral tissues. However, gel filtration of acid alcohol extracts of liver and kidney of animals injected 20 minutes earlier with labeled proinsulin does not disclose conversion to an insulin-like protein.

There is still insufficient information concerning the biological potency of proinsulin in the intact animal. Porcine proinsulin has an activity of 3 units per mg. in the mouse convulsion assay.[4, 5] We have observed rapid and profound hypoglycemia in alloxan-diabetic rats injected intravenously or subcutaneously with bovine proinsulin. Furthermore, proinsulin restores hepatic glucokinase activity and glycogen stores to normal levels in alloxan-diabetic rats.

CIRCULATING PROINSULIN

When samples of human serum and urine were extracted with acid ethanol and gel filtered on Sephadex G-50 in molar acetic acid or polyacrilamide (P-30) in 3 M acetic acid, two peaks of immunologically reactive material were found in many samples.[21] The earlier one corresponded in position to proinsulin and the second to insulin. In order to test whether proinsulin was converted to insulin or degraded during the clotting process, labeled and unlabeled proinsulin was added to freshly drawn blood. Eighty to 100 per cent of the proinsulin was recovered after extraction and column fractionation, and no material was found in the insulin peak. Furthermore, the total quantity of immunologically reacting material in column fractions after gel filtration was similar to that determined by direct assay of unextracted sera.

When the column fractions were assayed using an antiserum to porcine insulin, a tracer of ^{131}I-porcine insulin, and a human insulin standard, the early eluting material in normal subjects comprised 5 to

20 per cent of the insulin peak. The amount of proinsulin increased 2 to 4 hours after oral glucose. When fractions from the early peaks were combined and assayed in serial dilutions, the curve did not lie parallel to that of bovine or porcine proinsulin in assays using their specific antisera and tracers. The slope also differed from that of human insulin, but corresponded to a human proinsulin standard. Further characterization of this material must await the availability of human proinsulin antiserum and a more detailed examination of its properties.[24] At present the separation of proinsulin and insulin by gel filtration is a very time-consuming technique, and simpler methods, such as thin layer Sephadex chromatography, may greatly facilitate this approach.[26]

Other forms of immunologically reactive circulating insulin have been described. An abnormal insulin found in the serum of patients with juvenile diabetes was shown to be significantly more resistant to insulinase than endogenous insulin from healthy subjects.[9] Roth and coworkers[11, 20] have fractionated plasma samples on Sephadex columns and found two peaks of immunoreactive insulin. The second peak was coincident with pancreatic insulin ("little" insulin), while the first component was called "big" insulin. In five normal adults, "big" insulin 15 minutes after 100 gm. glucose ranged from 0 to 5 per cent, whereas at 2 hours it had risen to 5 to 28 per cent of the total immunoassayable insulin. Similar results were found after insulin secretion had been stimulated by arginine or tolbutamide.[11] The proportion of "big" insulin was markedly increased during the latter part of glucose tolerance tests performed on subjects who had been fasted for 72 hours. The ratio of "big" insulin to total insulin in three young, thin adults with glucose intolerance, acromegalics, and patients with myotonic dystrophy was similar to control values.[11] Obese subjects, on the other hand, tended to have higher levels of "big" insulin, particularly at 15 minutes. Although the relationship of proinsulin to these components is not yet clear, it seems likely that "big" insulin may largely consist of proinsulin or related proteins.

Studies concerning the effects of proteins which are immunologically related to insulin, but which cross react nonidentically in the insulin immunoassay system, have been carried out.[11, 16] On the basis of their experiments using mixtures of bovine and porcine insulins, Gorden and Roth[11] have suggested that in the assay of total plasma insulin on unfractionated diluted plasma, the presence of "big" insulin will produce minimal errors. Although the error would be magnified if the antibody used discriminated markedly between "big" and "little" insulin, or the quantity of "big" insulin increased, the discrepancy would still be of little physiological importance.[11] On the other hand, Lawrence and Kirsteins[16] have shown that serial dilution of certain sera does not follow the curve of a standard of porcine insulin closely, and that the measurement of total insulin will vary significantly depending upon the dilution chosen for assay.

Although we had hoped to be able to measure proinsulin directly in unextracted plasma with a specific antiserum which does not cross react with insulin,[37] this may not be feasible due to the presence of the

C-peptide in the circulation. This possibility has been confirmed in preliminary experiments on extracted fractionated bovine plasma, for displacement of ^{131}I-proinsulin from its specific antiserum occurred in the proinsulin region as well as in the area where both insulin and the C-peptide coelute.[23]

SUMMARY

The biosynthesis of insulin occurs in the beta cell from a single chain precursor protein, proinsulin. Proinsulin consists of the A and B chains of insulin linked by a polypeptide segment of approximately 30 amino acids. Conversion to insulin probably occurs in proximity to the Golgi apparatus while the protein is being packaged into secretory granules. The connecting peptide is then retained with insulin in the granule and is secreted together with it into the circulation.

Proinsulins have been isolated from first crystals of commercial insulin. The structures of porcine and bovine proinsulins are known and show considerable difference in the sequences of their connecting segments. This is also reflected in the distinctive immunological properties of the various species of proinsulin. The biological responses of many tissues to proinsulin is quantitatively similar to insulin, but proinsulin is much less active. The possible activation of proinsulin to an insulin-like molecule in adipose tissue and muscle has been considered.

Proinsulin and "big" insulin have been detected in the blood of healthy subjects. The levels of circulating proinsulin in different metabolic states and its physiological significance outside the pancreas are still unknown.

ADDENDUM

Since this manuscript was submitted for publication, new data on circulating proinsulin in humans have been obtained (Melani, F., Rubenstein, A. H., and Steiner, D. F.: Human serum proinsulin: Characterization and measurement in normal and obese subjects. J. Clin. Invest., in press). The serum proinsulin showed immunological identity with a human pancreatic proinsulin standard (Rubenstein, A. H., Melani, F., Pilkis, S., and Steiner, D. F.: Proinsulin: Secretion, metabolism, immunological and biological properties. Postgrad. Med., 45:475-581, 1969) and was able to be converted to an insulin-like component after exposure to trypsin. The absolute fasting values of proinsulin in normal subjects ranged between 0.05 and 0.4 ng. per ml. After oral administration of glucose, the proinsulin levels tended to rise similarly to those of insulin. Obese subjects with hyperinsulinemia had higher fasting levels of proinsulin and showed a greater rise following glucose. However, as these levels of proinsulin coexisted with raised insulin concentrations, the relative proportion of the two hormones was similar to that observed in the normal group.

REFERENCES

1. Bromer, W. W., and Patterson, J. M.: Refolding of reduced porcine proinsulin. Fed. Proc., 28:343, 1969.
2. Cerasi, E., and Luft, R.: "What is inherited—What is added"—Hypothesis for the pathogenesis of diabetes mellitus. Diabetes, *16*:615-627, 1967.

3. Challoner, D. R.: Differential effects of porcine proinsulin on rat fat pieces and fat cells. Proc. Amer. Soc. Clin. Invest., 61st Annual Meeting, 1969.
4. Chance, R. E., and Ellis, R. M.: Proinsulin: single chain precursor of insulin. Arch. Int. Med., *123*:223-236, 1969.
5. Chance, R. E., Ellis, R. M., and Bromer, W. W.: Porcine proinsulin: Characterization and amino acid sequence. Science, *161*:165-167, 1968.
6. Clark, J. L., Cho, S., Rubenstein, A. H., and Steiner, D. F.: Isolation of a proinsulin connecting peptide fragment (C-peptide) from bovine and human pancreas. Biochem. Biophys. Res. Commun., *35*:456-461, 1969.
7. Clark, J. L., and Steiner, D. F.: Studies on insulin biosynthesis in the rat and the demonstration of two proinsulins. Proc. Nat. Acad. Sci., *62*:278-285, 1969.
8. Colwell, J. A., and Lein, A.: Diminished insulin response to hyperglycemia in prediabetes and diabetes. Diabetes, *16*:560-565, 1967.
9. Elliott, R. B., O'Brien, D., and Roy, C. C.: An abnormal insulin in juvenile diabetes mellitus. Diabetes, *14*:780-787, 1965.
10. Gliemann, J., personal communication, 1968.
11. Gorden, P., and Roth, J.: Circulating insulins: "Big" and "little." Arch. Int. Med., *123*: 237-247, 1969.
12. Grant, P. T., and Reid, K. B. M.: Biosynthesis of an insulin precursor by islet tissue of cod (Gradus callarias). Biochem. J., *110*:281-288, 1968.
13. Jamieson, J. D., and Palade, G. E.: Intracellular transport of secretory proteins in the pancreatic exocrine cell. J. Cell Biol., *34*:577-615, 1967.
14. Kipnis, D. M.: Insulin secretion in diabetes mellitus. Ann. Int. Med., *69*:891-901, 1968.
15. Lacy, P. E., personal communication, 1969.
16. Lawrence, A. M., and Kirsteins, L.: The effect of proinsulin on the radioimmunoassay of insulin and its possible relation to some instances of hyperinsulinemia. J. Lab. Clin. Med., 72:988, 1968.
17. Narahara, N. T., personal communication, 1968.
18. Nolan, C., Margoliash, E., and Steiner, D. F. (in preparation).
19. Pilkis, S., Rubenstein, A. H., and Steiner, D. F.: The biological activity of bovine proinsulin (in preparation).
20. Roth, J., Gorden, P., and Pastan, I.: "Big" insulin: A new component of plasma insulin detected by immunoassay. Proc. Nat. Acad. Sci., *61*:138-145, 1968.
21. Rubenstein, A. H., Cho, S., and Steiner, D. F.: Evidence for proinsulin in human urine and serum. Lancet, *1*:1353-1355, 1968.
22. Rubenstein, A. H., Cho, S., and Steiner, D. F.: Comparative immunologic properties of bovine, porcine and human proinsulin. J. Lab. Clin. Med., 72:1010, 1968.
23. Rubenstein, A. H., Clark, J. L., Melani, F., and Steiner, D. F.: Secretion and circulation of the proinsulin C-peptide. Nature, in press.
24. Rubenstein, A. H., and Steiner, D. F.: Human proinsulin: Some considerations in the development of a specific immunoassay. Proceedings, International Symposium on Early Diabetes, Marbella, Spain. New York, Academic Press, to be published.
25. Rubenstein, A. H., Steiner, D. F., Cho, S., Lawrence, A. M., and Kirsteins, L.: Immunological properties of bovine proinsulin and related fractions. Diabetes, *18*:598-605, 1969.
26. Ryan, W. G., and Robbins, P.: Thin layer Sephadex chromatography of insulin and proinsulin. Clin. Res., *17*:394, 1969.
27. Schmidt, V. D. D., and Arens, A.: Proinsulin von Rind: Isolierung, Eigenschaften und seine Aktivierung durch Trypsin. Hoppe Seyler Z. Physiol. Chem., *349*:1157-1168, 1968.
28. Shaw, W. N., and Chance, R. E.: The effect of porcine proinsulin in vitro on adipose tissue and diaphragm of the normal rat. Diabetes, *17*:737-745, 1968.
29. Steiner, D. F.: Evidence for a precursor in the biosynthesis of insulin. Trans. N.Y. Acad. Sci., *30*:60-68, 1967.
30. Steiner, D. F., and Clark, J. L.: The spontaneous reoxidation of reduced beef and rat proinsulins. Proc. Nat. Acad. Sci., *60*:622-629, 1968.
31. Steiner, D. F., Clark, J. L., Nolan, C., Rubenstein, A. H., Margoliash, E., Aten, B., and Oyer, P. E.: Proinsulin and the biosynthesis of insulin. Recent Progress Hormone Res., 25:207-282, 1969.
32. Steiner, D. F., Cunningham, D., Spigelman, L., and Aten, B.: Insulin biosynthesis: Evidence for a precursor. Science, *157*:697-700, 1967.
33. Steiner, D. F., Hallund, O., Rubenstein, A. H., Cho, S., and Bayliss, C.: Isolation and properties of proinsulin, intermediate forms, and other minor components from crystalline bovine insulin. Diabetes, *17*:725-736, 1968.

34. Steiner, D. F., and Oyer, P. E.: The biosynthesis of insulin and a probable precursor of insulin by a human islet cell adenoma. Proc. Nat. Acad. Sci., 57:473-480, 1967.
35. Tung, A. K., and Yip, C. C.: The biosynthesis of insulin and "proinsulin" in foetal bovine pancreas. Diabetologia, 4:68-70, 1968.
36. Wrenshall, G. A., Bogoch, A., and Ritchie, R. C.: Extractable insulin of pancreas. Diabetes, 1:87-105, 1952.
37. Yip, C. C., and Logothetopoulos, J.: A specific anti-proinsulin serum and the presence of proinsulin in calf serum. Proc. Nat. Acad. Sci., 62:415-419, 1969.

950 East 59th Street
Chicago, Illinois 60637

The Effect of Insulin on Protein Synthesis in Muscle

An Example of an Approach to the Mechanism of Hormone Action

JAMES J. CASTLES, M.D.*

Insulin has a multitude of effects on the metabolism of muscle. These include stimulation of sugar and amino acid transport and an increase in glycogen, lipid, protein, and nucleic acid synthesis.[16] The legion of effects mediated by insulin may be due to multiple primary effects of the hormone. However, most investigators concerned with mechanisms of hormone action prefer to believe that a hormone ultimately has a single or unitary action, and that all of those effects now recognized represent consequences of that action. Accepting the dictum that insulin has a unitary action, Wool has hypothesized that the effects of insulin on muscle are secondary to the ability of the hormone to stimulate protein synthesis.[16] This review will summarize some of the experiments which attempt to locate the ultimate molecular site at which insulin stimulates protein synthesis in muscle; a detailed review of some of this work is available.[24]

It should be recognized that numerous hypotheses on the mode of insulin action are available. Some of the alternative hypotheses can be found in reviews by Levine and Goldstein,[7] Tepperman and Tepperman,[14] and Krahl.[5]

The Effect of Insulin on Protein Synthesis in Muscle is Independent of Substrate Transport

The experiments of Levine and colleagues[7] documented that insulin stimulated the transport of monosaccharides in muscle. It had also been observed that incorporation of amino acids into muscle proteins was increased by insulin.[4, 13] However, Levine and Goldstein[7] postulated that the ability of insulin to stimulate protein synthesis (and many of the other effects of insulin) was secondary to the primary effect of the hor-

*Assistant Professor, Department of Medicine, University of Chicago School of Medicine, Chicago, Illinois

mone on sugar transport. This idea was tested using isolated rat diaphragm; despite the absence of glucose or other sugars in the incubation medium, insulin still increased the incorporation of amino acids into protein.[9, 18] Clearly, the effect of insulin on protein synthesis did not require the concomitant stimulation of sugar transport.

Insulin might increase protein synthesis by stimulation of amino acid transport into muscle cells. Indeed, insulin does accelerate the transport of amino acids into muscle.[2, 3] However, under circumstances where amino acid transport cannot be involved, insulin still stimulates protein synthesis.[17]

It would appear that the effect of insulin on protein synthesis is through an effect on some intracellular process. Prior to considering the intracellular site of insulin action, a brief review of the present state of our knowledge of the mechanism of protein synthesis seems in order.

Current Scheme of Protein Biosynthesis

It is well known that the ultimate determinant of the proteins that a cell may possess is the genetic material, DNA. In order to synthesize the protein coded for by a gene, the initial step is the synthesis of an RNA molecule complementary to the DNA sequence of the gene; this is messenger RNA (mRNA). The sequences of triplets of nucleic acid bases (codons) in mRNA will specify the sequence of amino acids in the protein to be made and, thus, the specific protein. In mammalian cells, mRNA must be transported from its site of synthesis, the nucleus, to the cytoplasm; this process is poorly understood. The translation of mRNA into protein is a complex but precise process which requires (1) energy; (2) the appropriate ionic environment; (3) a population of RNA molecules which are capable of recognizing the mRNA codons which specify the amino acid that is covalently bound to the RNA species (transfer RNA or tRNA); (4) a ribosome, an intracellular organelle consisting of a small subunit (sedimentation coefficient (S) of 40) and a large subunit (60S), each with its own RNA and specific complement of proteins; and (5) numerous cytoplasmic proteins which are requisite in each of the steps outlined below.

The steps involved in the translation of mRNA can conveniently be divided into the following scheme (Fig. 1):

1. Binding of mRNA by ribosomes.
2. Initiation of translation of mRNA. A precise initiation signal must exist in each mRNA in order to insure that the protein coded for will be synthesized completely and that the sequence of codons will be read in proper phase. In bacteria, a specific codon signals initiation by stimulating binding of a specific tRNA (formylmethionyl-tRNA) to the initiator codon-ribosome complex.
3. Recognition of the second codon in the message by its specific tRNA and binding of this tRNA.
4. Peptide bond formation between the initiating amino acid (amino terminus of the protein) and the next amino acid in sequence.
5. Movement of the ribosome (translocation) along the mRNA so the third codon is positioned for recognition by its tRNA.
6. Repetition of steps 3 to 5; until
7. A codon for terminating reading of the message and signaling completion of the protein is reached.

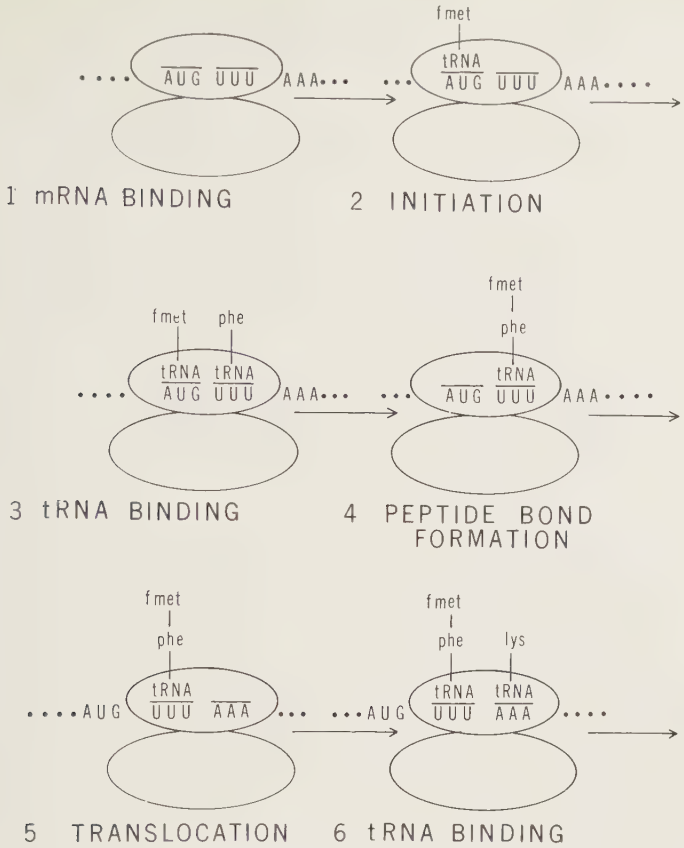


Figure 1. Scheme of protein biosynthesis. The details of each step can be found in the text.... AUGUUUAAA ... is a hypothetical messenger RNA molecule, where A is the nucleic acid base adenine, U is uridine, and G is guanine. The triplet AUG is the code word (codon) for a transfer RNA which has the amino acid, N-formylmethionine (f met), covalently attached to it. Similarly, UUU is the codon for phenylalanyl (phe)-tRNA and AAA is the codon for lysyl (lys)-tRNA.

There is reason to suspect that regulation of protein synthesis can occur at any of the steps outlined above.

The Intracellular Locus at which Insulin Stimulates Protein Synthesis

Insulin could increase protein synthesis by increasing the synthesis of RNA, specifically mRNA. It has been documented that insulin does increase RNA synthesis in muscle.[15] However, when muscle RNA synthesis is inhibited by actinomycin D, insulin is still capable of increasing protein synthesis.[19] The results suggest that insulin acts at a site distal to the synthesis of mRNA.

At this stage, the observation that a cell-free system from skeletal muscle of alloxan-diabetic rats was less capable of synthesizing proteins than the same system from normal rats greatly facilitated subsequent

analysis.[21] The defect in diabetic animals was found to reside in the ribosomes and not in the cytoplasmic protein factors; it could be corrected by administering insulin to diabetic animals prior to isolating ribosomes. In fact, as little as 0.1 unit of insulin was sufficient to increase protein synthesis by diabetic muscle ribosomes 5 minutes after administration. The same amount of insulin produced no detectable change in blood glucose concentration in this same time interval. However, insulin added in vitro to muscle ribosomes (normal or diabetic) did not affect their ability to synthesize protein.[21]

It is important to note that the above experiments were performed in the absence of added mRNA; that is, all in vitro protein synthesis was directed by "endogenous" mRNA which remained bound to the ribosomes during their preparation. Consequently, the defect present in diabetes and corrected by insulin may be in the number of mRNA molecules bound to ribosomes. Indeed, direct measurement of active ribosomes (those containing mRNA in the process of translation) revealed fewer active ribosomes from diabetic rat muscle than from normal.[22]

A closer look was taken at the events occurring between administration of insulin to a diabetic animal and isolation of the ribosomes. When RNA synthesis was inhibited in diabetic rats prior to giving them insulin, the hormone still corrected the defect in the ribosomes (Table 1). However, when diabetic animals were given an inhibitor of protein synthesis prior to receiving insulin, the defect in the ribosomes could not be corrected (Table 2). These experiments suggest that mRNA is synthesized or already present in a normal quantity in diabetic animals. However, diabetic ribosomes lack a factor necessary for the binding or translation of mRNA; insulin stimulates the synthesis of this factor ("translation factor"), which is probably a protein.[20]

The "translation factor" may be a cytoplasmic protein which modifies the function of ribosomes or it may be a part of the ribosome itself, i.e., a ribosomal protein. Analysis of incorporation of radioactive amino acids into proteins by discontinuous gel electrophoresis revealed that there was a disproportionate synthesis of at least one protein in both the cytoplasm and the ribosomes after the administration of insulin to diabetic animals.[6]

Table 1. *Protein Synthesis by Ribosomes from Skeletal Muscle: Effect of Diabetes and Insulin in the Presence of Actinomycin D*

SOURCE OF RIBOSOMES	NO INHIBITOR	ACTINOMYCIN D
Normal	1418	1525
Diabetic	718	640
Diabetic + Insulin	1100	1167

Actinomycin D or saline was given to the animals 75 minutes prior to killing; in the instances where insulin was also administered, this was given 15 minutes after the actinomycin or saline.

The results are expressed as counts per minute of ^{14}C-phenylalanine incorporated into protein by 130 μg. of ribosomes. The data are from Wool and Cavicchi.[20]

Table 2. *Protein Synthesis by Ribosomes from Skeletal Muscle: Effect of Diabetes and Insulin in the Presence of Puromycin*

SOURCE OF RIBOSOMES	NO INHIBITOR	PUROMYCIN
Normal	674	554
Diabetic	237	248
Diabetic + Insulin	698	320

The animals were given puromycin or saline 135 minutes prior to being killed; the rats given insulin received it 1 hour prior to death.

The results are expressed as cpm of ^{14}C-phenylalanine incorporated into protein by 130 μg. of ribosomes. The data are from Wool and Cavicchi.[20]

Further understanding of the role of the "translation factor" was facilitated by isolating active subunits from muscle ribosomes of normal and diabetic rats.[10] By hybridization of normal and diabetic ribosomal subunits (combining a subunit derived from a normal animal with a complementary subunit from a diabetic), Martin and Wool determined that the functional defect of the ribosomes from the diabetic animals was a property of the larger (60S) ribosomal subunit (Table 3). It was correctable by the administration of insulin prior to the isolation of the ribosomes. Significantly this defect was detectable using an exogenous mRNA; this is further evidence against the possibility that mRNA synthesis or degradation is involved in the action of insulin on protein biosynthesis. In addition, it has been documented that some functions of the smaller subunit (40S) are normal in diabetic ribosomes; these functions are mRNA binding[11] and mRNA-directed tRNA binding to the 40S subunit.[1] However, mRNA-directed tRNA binding to diabetic 80S ribosomes demonstrates a defect similar to that seen in protein synthesis, and this is also a property of the 60S subunit.[12]

To date, no physical or chemical change has been noted in diabetic muscle ribosomes or subunits. This is true for ribosomal RNAs[23] and ribosomal proteins.[8] Because of the limitations of the present analytical techniques, it is still possible that ribosomes from diabetic animals do lack a protein whose synthesis is controlled by insulin. Alternatively, the functional defect in the diabetic ribosome may be due to a conformational change of a ribosomal protein. An additional possibility is that

Table 3. *Protein Synthesis by Hybrids of Ribosomal Subunits from Muscle of Normal and Diabetic Rats**

RIBOSOMAL PARTICLES	cpm ^{3}H-PHENYLALANINE INCORPORATED INTO PROTEIN BY 20 μg. RIBOSOMES
Normal 40S + Normal 60S	1611
Diabetic 40S + Diabetic 60S	816
Normal 40S + Diabetic 60S	634
Diabetic 40S + Normal 60S	1722

*The data are from Martin and Wool.[10]

insulin controls the synthesis or conformation of a labile cytoplasmic protein which is requisite for activating the 60S subunit but is not a component of the particles.

SUMMARY

Insulin stimulates protein synthesis in muscle at an intracellular locus. The experiments suggest that insulin either alters the function of a labile protein ("translation factor") or induces the formation of that protein by initiating the translation of its stable mRNA. The "translation factor" is either a component of a larger ribosomal subunit or controls the function of this subunit without necessarily becoming a component of the particle.

ACKNOWLEDGMENTS

I am indebted to Professor Ira G. Wool for his criticisms of the manuscript and for permission to cite some experiments prior to their publication.

REFERENCES

1. Castles, J. J., unpublished experiments.
2. Castles, J. J., and Wool, I. G.: Effect of puromycin and insulin on amino acid accumulation by isolated intact rat diaphragm. Biochem. J., *91*:11C, 1964.
3. Kipnis, D. M., and Noall, M. W.: Stimulation of amino acid transport by insulin in the isolated rat diaphragm. Biochim. Biophys. Acta, *28*:226, 1958.
4. Krahl, M. E.: Incorporation of C^{14} amino acids into peptides by normal and diabetic rat tissues. Science, *116*:524, 1952.
5. Krahl, M. E.: Speculation on the action of insulin, with a note on other hypoglycemic agents. Persp. Biol. Med., *1*:69, 1957.
6. Kurihara, K., and Wool, I. G.: Effect of insulin on the synthesis of sarcoplasmic and ribosomal proteins of muscle. Nature, *219*:721, 1968.
7. Levine, R., and Goldstein, M.: On the mechanism of action of insulin. Recent Prog. Hormone Res., *11*:343, 1955.
8. Low, R., and Wool, I. G.: Skeletal muscle ribosomal proteins: General characteristics and effect of diabetes. Biochim. Biophys. Acta, in press.
9. Manchester, K. L., and Young, F. G.: Effect of insulin on incorporation of amino acids into protein of normal rat diaphragm in vitro. Biochem. J., *70*:353, 1958.
10. Martin, T. E., and Wool, I. G.: Formation of active hybrids from subunits of muscle ribosomes from normal and diabetic rats. Proc. Nat. Acad. Sci., *60*:569, 1968.
11. Rolleston, F. S., Martin, T. E., and Wool, I. G.: The binding and translation of TYMV RNA by skeletal muscle ribosomes from normal and diabetic rats. In preparation.
12. Rolleston, F. S., Martin, T. E., and Wool, I. G.: The effects of magnesium on polyuridylic acid directed protein synthesis and on binding of phenylalanyl-transfer RNA by ribosomes from muscle of normal and diabetic rats. In preparation.
13. Sinex, F. M., MacMullen, J., and Hastings, A. B.: The effect of insulin on the incorporation of C^{14} into the protein of rat diaphragm. J. Biol. Chem., *198*:615, 1952.
14. Tepperman, J., and Tepperman, H. M.: Some effects of hormones on cells and cell constituents. Pharmacol. Rev., *12*:301, 1960.
15. Wool, I. G.: Effect of insulin on nucleic acid synthesis in isolated rat diaphragm. Biochim. Biophys. Acta, *68*:28, 1963.
16. Wool, I. G.: Insulin and protein biosynthesis. *In* Litwack, G., and Kritchevsky, D., eds.: Action of Hormones in Molecular Processes. New York, John Wiley & Sons, 1964.
17. Wool, I. G.: Relation of effects of insulin on amino acid transport and on protein synthesis. Fed. Proc., *24*:1060, 1965.
18. Wool, I. G., and Krahl, M. E.: Incorporation of C^{14}-amino acids into protein of isolated diaphragm: An effect of insulin independent of glucose entry. Amer. J. Physiol., *196*: 961, 1959.

19. Wool, I. G., and Moyer, A. N.: Effect of actinomycin and insulin on the metabolism of isolated rat diaphragm. Biochim. Biophys. Acta, *91*:248, 1964.
20. Wool, I. G., and Cavicchi, P.: Insulin regulation of protein synthesis by muscle ribosomes: Effect of the hormone on translation of messenger RNA for a regulatory protein. Proc. Nat. Acad. Sci., *56*:991, 1966.
21. Wool, I. G., and Cavicchi, P.: Protein synthesis by skeletal muscle ribosomes. Effect of diabetes and insulin. Biochemistry, *6*:1231, 1967.
22. Wool, I. G., and Kurihara, K.: Determination of the number of active muscle ribosomes: Effect of diabetes and insulin. Proc. Nat. Acad. Sci., *58*:2401, 1967.
23. Wool, I. G., Stirewalt, W. S., and Moyer, A. N.: Effect of diabetes and insulin on nucleic acid metabolism of heart muscle. Amer. J. Physiol., *214*:825, 1968.
24. Wool, I. G., Stirewalt, W. S., Kurihara, K., Low, R. B., Bailey, P., and Oyer, D.: Mode of action of insulin in the regulation of protein biosynthesis in muscle. Recent Prog. Hormone Res., *24*:139, 1968.

950 East 59th Street
Chicago, Illinois 60637

V. NEW TRENDS IN CONCEPTS AND THERAPY OF HEART DISEASE

Introduction

There have been few fields in medicine influenced more significantly by the advancing knowledge in cellular biology than the cardiovascular area—ranging from the still unclear relationship of cardiac hypertrophy to intracellular protein synthesis, to heart failure and its basis in cellular bioenergetics of the contractile mechanisms, and to the clinical concern with the link between cellular excitation and muscular contraction, or to the nature of ionic transport across cell membranes and its influence by therapeutic agents. These concepts and techniques developed for skeletal and cardiac muscle are being extended to involve and interpret the mechanisms of peripheral vascular control.

Two reports, one on the nature of cardiac muscle hypertrophy closely following the thoughts on hormone action in protein synthesis expressed in the last article, and one dealing with cellular action of cardiac glycosides, are given as examples of how clinical cardiology is likely to tap the basic resources of the biologic sciences as a continuing source of new stimulation. The field is advancing rapidly on all fronts, however, and two papers, one on modern management of the "pump failure" in acute myocardial infarction, and one on the controversial use of "paired" external stimulation to suppress abnormal tachyarrhythmias, illustrate how many eminently practical situations still pose significant investigational problems. The section—and the issue—closes with a new look at so-called "functional" apical murmurs masking a great variety of pathological lesions almost unheard of a few years ago. More than any other, this section seems to open the door wide to the future.

H. H. H.

Cellular and Subcellular Basis of Cardiac Hypertrophy

MURRAY RABINOWITZ, M.D.*

K. G. NAIR, M.D., Ph.D.**

RADOVAN ZAK, Ph.D.***

The consequences and the mechanisms of cardiac hypertrophy have been the subject of intense study by clinicians and investigators for many decades. The development of ventricular hypertrophy appears to be a fundamental adaptive mechanism of the heart in response to an increased work load. The increased hemodynamic burden on the heart may be either a pressure overload, as in aortic stenosis, pulmonic stenosis, systemic hypertension, or pulmonary hypertension, or a volume overload, such as in insufficiency of the atrioventricular or semilunar valves. Cardiac muscle enlargement may also be secondary to high cardiac output states such as anemia, arteriovenous fistula, intracardiac shunts or hypoxia. Finally, ventricular hypertrophy may also be present in patients with congestive cardiac failure secondary to ischemic heart disease or cardiomyopathy.

Although cardiac hypertrophy is initially a favorable response to an increased workload, there may be a limit to the usefulness of such an adaptive process. When the hemodynamic burden is present for a protracted time, the hypertrophic response may be both ineffective and inappropriate to the situation; or, with acutely imposed overload, such as occurs with a ruptured papillary muscle, there may not be time enough for an adequate hypertrophic response.

From the Department of Medicine (Cardiology) and the Argonne Cancer Research Hospital (operated by the University of Chicago for the United States Atomic Energy Commission), University of Chicago, Chicago, Illinois

*Professor of Medicine and Biochemistry

**Formerly Assistant Professor of Medicine and Physiology; recipient of U.S. Public Health Service Career Development Award No. K4-HE-38898; presently Associate Professor of Medicine, University of Pennsylvania School of Medicine, Philadelphia, Pennsylvania

***Assistant Professor of Medicine; Research Associate, Department of Biochemistry

This investigation was supported in part by grants from the National Heart Institute, National Institutes of Health (HE-09172, HE-04442, and HE-05447), from the Chicago Heart Association; and from the Myocardial Infarction Research Unit (contract PH-4368133).

Clinically the onset of hypertrophy may be taken as an unfavorable event indicative of severe disease. The hypertrophic process in most cases, however, would appear to be compensatory, one which enables the heart to perform a larger work load and still remain compensated. Congestive failure, however, may result if the hypertrophy is inadequate. Several variables may affect the clinical course and determine whether failure occurs immediately or is delayed. The rapidity of onset of the stimulus for hypertrophy may be an important determinant. Acute onset of systemic or pulmonary hypertension is more likely to lead to failure than is gradual appearance. The age of onset of hypertrophy may also be of great importance. Patients with congenital pulmonic stenosis with right ventricular systolic pressures well over 100 mm. Hg may survive for decades with no evidence of failure, whereas right heart failure may appear in cor pulmonale with pulmonary artery pressures that are considerably lower. The state of the myocardium which is called to do the extra work is of course also of prime importance. Hearts of elderly peopie with ischemic heart disease are more likely to fail under stress than those with a normal coronary circulation.

Although these and other clinical observations are well recognized, the exact mechanisms by which cardiac hypertrophy develops are not known. The great increase in our understanding of the mechanism and control of nucleic acid and protein synthesis now enables us to reinvestigate this problem at the cellular and subcellular level: we have begun to unravel some of the intimate biochemical and cytological processes which occur in cardiac hypertrophy.

HYPERTROPHY, HYPERPLASIA, AND DNA SYNTHESIS

Growth of a differentiated tissue such as heart muscle is thought to occur by enlargement of individual cells rather than by proliferation and increase in cell number. The process of muscle cell enlargement without proliferation could lead to conditions that result in impairment of function. Pathologists have long considered that if the vascular supply of the heart does not increase commensurately with the increase in cardiac mass the result would be relative ischemia.[38] Recent studies, however, have cast doubt on this hypothesis.[21] Another possibility to be considered is an insufficient synthesis of cellular organelles such as mitochondria and enzyme systems in relation to increase in cell size. Specific metabolic process may thereby be limited. Still another suggestion is that the increase in cell size results in a decrease in the ratio of nucleus to cytoplasm.[27] The DNA in the cell may thereby become insufficient to provide adequate information for the synthesis of all necessary proteins.

The question of whether cardiac enlargement secondary to increased workload indeed is limited to an increase in cell size has recently been evaluated. In experimental animals, cardiac enlargement can be produced by a variety of means, such as surgical banding of the ascending aorta[30] or of the pulmonary artery in rats,[3] rabbits,[31] and dogs,[25] by the production of systemic renal hypertension,[6] or by anemia[45] or hypoxia.[2]

Constriction of the ascending aorta of the rat with a silver band results in an increase in heart weight and in RNA and DNA content of 20 to 50 per cent within 48 hours of surgery.[14, 19, 30] The increase in total DNA is of particular significance to our definition of hypertrophy. If cell size alone were increasing without cell proliferation, no increase in DNA should have been observed, because the DNA content in each cell with a nucleus containing the normal two sets of chromosomes is constant. In cases where there is no increase in the number of sets of chromosomes per nucleus, a DNA increment can only mean that nuclear division and probably cell division have occurred.

The paradox of increased cardiac DNA in a tissue presumably growing by hypertrophy alone has been investigated by several laboratories. It was established that in the experimental animal the increase in DNA could not be accounted for by polyploidy (multiplication of chromosome number not followed by nuclear or cell division).[13, 14, 40] Polyploid frequency did increase from 2 per cent in sham-operated control hearts to over 4 per cent in hypertrophied hearts, but this increase was not nearly enough to account for the 30 to 50 per cent increase in cardiac DNA content. Radioautographic studies showed that radioactive thymidine was incorporated almost exclusively into nuclei of connective tissue cells. Almost no synthesis of DNA in muscle cells could be detected.[6, 15, 26, 28] An increase in the mitotic index was also observed in hypertrophied hearts, but mitoses were limited to connective tissue cells.[14] Quantitative differential nuclear counts also showed that the total number of muscle nuclei did not increase, whereas the total number of connective tissue nuclei did increase. There was also no evidence for infiltration of inflammatory cells in hypertrophy.

Therefore it could be concluded that, at least in the model of acutely induced cardiac enlargement, no significant division of muscle cell occurs, but that connective tissue hyperplasia is striking. An increase in cardiac collagen content accompanies the increase in number of connective tissue cells.[2, 5] In contrast to the absence of division of muscle cells in adult hearts, increases in muscle cells are observed with aging in young and embryonic hearts. This increase appears to be secondary to a proliferation of undifferentiated precursor cells, rather than to division of differentiated muscle cells.[18]

Studies of developing skeletal and cardiac muscle in embryonic tissue and in tissue culture have shown that mitoses are almost exclusively limited to undifferentiated primitive cells, the myoblasts.[44] When the myoblast differentiates into a mature cell containing myofibrils, the cell's ability to divide is lost immediately or soon thereafter.[44] In the case of skeletal muscle the myoblasts, perhaps in the form of the "satellite cells" (subsarcolemmal cells containing a large nucleus and very little cytoplasm), persist in the adult tissue and may resume mitosis after injury.[35] Skeletal muscle regeneration may develop into a very active process.[20] It is unclear whether primitive myoblasts also persist in adult cardiac tissue.[43] Clinically it is apparent that following a myocardial infarction regeneration of necrotic muscle cells does not occur. Experimental evidence for regeneration of adult cardiac muscle cells has never been obtained.[37] Therefore if myoblasts are retained in

mature cardiac muscle, even extreme pathological stimuli are inadequate to reinstate muscle cell division.

In embryonic cardiac tissue and in the hearts of young organisms, however, myoblasts are probably present and may proliferate.[12] Production of cardiac enlargement at this time by congenital lesions may theoretically result in a true muscle hyperplasia as well as hypertrophy. The more efficient process of increasing the number of cardiac muscle cells, as well as increasing their size, may account for examples of severe prolonged hypertrophy without congestive failure that occur in congenital pulmonic and aortic stenosis, and in pulmonary hypertension associated with congenital heart disease.

The differentiated muscle cell appears to be incapable of dividing under ordinary, normal, and pathological conditions. Of interest, however, is the observation that differentiated muscle cells in tissue culture may be stimulated to resume proliferation by infection with the oncogenic polyoma virus.[46] The inherent capacity of the cell for cell division therefore has not been lost. Normal control mechanisms are merely inadequate to induce mitosis. It remains possible that some sort of artificial stimuli may yet be discovered which could be effective in stimulating cardiac muscle regeneration after muscle death in myocardial infarction. This is probably no more remote than the suggestion that it may ultimately be possible to alter a patient's genetic material and thereby modify genetic disease by transduction or transformation.

PROTEIN AND RNA SYNTHESIS

Cardiac tissue is composed of a large variety of cells. In addition to contractile muscle cells, also present are Purkinje cells, connective tissue, vascular elements, and inflammatory cells. During hypertrophy all elements increase in size or number to a varying degree. Since one of the most important functional components of all cells is protein, attention has been directed to the changes in synthesis of cardiac protein in hypertrophy. In experimental cardiac hypertrophy there is a marked stimulation of synthesis of both RNA and protein in the heart.[8, 19] The increase in cardiac RNA and protein content parallels the increase in heart weight in the early stages of the process.[30] Experiments using incorporation of radioactive amino acids into protein, or labeled nucleosides into RNA, indicate that the increment in RNA and protein content is largely secondary to an increased rate of synthesis. However, we now know that changes in the rate of protein and RNA turnover may also contribute to their accumulation; such a process has been demonstrated for skeletal muscle protein during induced hypertrophy.[11]

It is not surprising that RNA as well as protein increases in cardiac hypertrophy, since it is well recognized that at least three types of RNA are involved in the process of protein synthesis—ribosomal, messenger, and transfer RNA. Quantitatively, ribosomal RNA comprises 80 to 85 per cent of total cellular RNA, transfer RNA 10 to 20 per cent, and messenger RNA 1 to 2 per cent. The increase in RNA content in hyper-

trophy largely consists of ribosomal RNA, although a proportional increase in transfer and messenger RNA probably also occurs.

The assembly of proteins takes place on ribosomes, or rather on aggregates of ribosomes (polysomes) held together by messenger RNA. Ribosomes are composed of approximately 60 per cent ribosomal RNA (28S and 18S) and 40 per cent ribosomal proteins (20 to 30 varieties). Messenger RNA specifically codes for a complementary triplet of bases present in a transfer RNA. Since each transfer RNA carries only one specific amino acid, a messenger RNA effectively is coding for one amino acid. The three types of RNA involved in protein synthesis are all transcribed from complementary base sequences present in DNA by the enzyme RNA polymerase. Ordinarily most of the nuclear DNA is repressed and is not actively being utilized as a template for RNA synthesis. Various physiological or biochemical stimuli may derepress the DNA and thereby stimulate the synthesis of ribosomal RNA, or of messenger RNA coding for specific cell proteins.

An increase in messenger RNA synthesis without an accompanying increase in the number of ribosomes is the important mechanism used by microorganisms to increase the synthesis of specific proteins. Ribosomes are presumably present in excess. A coordinated synthesis of messenger RNA acting on existing ribosomes may direct the production of a series of proteins. This mechanism appears to be of secondary importance in experimental cardiac hypertrophy. Instead, new ribosomes seem to be made in order to synthesize more proteins. The large increase in cardiac RNA in hypertrophy is predominantly due to ribosomal RNA.[8, 19]

Ribosomes isolated from hypertrophied hearts are no more active in protein synthesis than ribosomes from control hearts.[29] Increased activity would be expected if more messenger RNA had been synthesized and had stimulated existing ribosomes. Furthermore, no detectable change in the pattern of polysomes is seen during density gradient centrifugation in cardiac hypertrophy.[29] A greater proportion of large polysomal aggregates would be expected if messenger RNA synthesis had stimulated existing ribosomes to synthesize protein. The principal mechanism accounting for the increase in protein synthesis therefore appears to be an increase in the number of ribosomes, probably associated with a proportional increase in messenger RNA. Only in very acute experiments using isolated perfused rat hearts has a change in polysome pattern, suggesting predominantly messenger RNA synthesis, been observed following elevation of aortic pressure.[41]

It is apparent from the above consideration that increased synthesis of RNA probably precedes and accompanies the increased protein synthesis in cardiac hypertrophy. Estimation of the rate of ribosomal RNA synthesis using radioactive uridine showed a four to fivefold increase 24 hours after aortic banding and a twofold increase 6 days after surgery.[19] The enzyme RNA polymerase, which synthesizes cellular RNA, is localized in the cell nucleus, and uses chromosomal DNA as a template to direct the synthesis of specific RNA. In an effort to come a step closer to the stimulus initiating cardiac hypertrophy, Nair et al. isolated nuclei from hypertrophied and sham-operated rats and compared

their RNA polymerase activity by a variety of assays.[30] Nuclear RNA polymerase was elevated within 24 hours after the induction of hypertrophy and remained elevated, but to a smaller extent, for 10 days. Analysis of the characteristics of the RNA synthesized by the nuclei, using nearest neighbor analysis, showed it to be similar to ribosomal RNA. Ribosomal RNA is also the predominant species of RNA synthesized at this time in the intact animal.

Thus one of the earliest detectable biochemical changes in cardiac hypertrophy is an increase in RNA polymerase activity. The mechanism of the increase is still unclear; it may be due to increased template activity of derepressed chromosomal DNA, to synthesis of new enzyme, or to release of inhibition or activation of the existing enzyme.

MECHANISM OF INITIATION OF CARDIAC HYPERTROPHY

Although increased RNA synthesis may be one of the earliest biochemical events leading to cardiac hypertrophy, the exact mechanism by which elevated pressure or volume overload on the heart produces the varied but integrated biochemical and cellular responses that comprise cardiac hypertrophy is still not known. Mechanisms similar to those originally postulated by Jacob and Monod[17] involving a repressor-derepressor system have been proved to exist in microorganisms, and probably also participate in the genesis of hypertrophy. Most cellular DNA normally is in a repressed state, incapable of acting as a template for RNA transcription. Physiological and ultimately biochemical changes can release the repression and thereby induce RNA synthesis. The repressor-derepressor mechanism has been demonstrated in E. coli[10, 32] but has not yet been conclusively shown to be present in higher organisms.

Even if we accept the hypothesis of a repressor-derepressor mechanism, the mechanism by which the physiological stimulus is converted into a biochemical repressor or derepressor is unknown. One possibility is that stretch of muscle, or of muscle membrane, may be the initial physiological stimulus. Under conditions of increased pressure or volume work, or secondary to cardiac dilatation, increased stretch of muscle fibers, or of associated membrane systems, may in some way initiate biochemical changes which act in a repressor system. How such a stretch stimulus can be converted to a biochemical change is not clear. The participation of such a mechanism is suggested by observations on the hypertrophy resulting from denervation of a hemidiaphragm in the rat.[9, 47] The denervated diaphragm hypertrophies only if the opposite hemidiaphragm is functioning. The stimulus for hypertrophy of the denervated diaphragm appears to be provided by its passive stretch by the functioning diaphragm.

Another possible mechanism by which increased work load may induce hypertrophy is that local tissue hypoxia or ischemia may result in accumulation of metabolites which act as biological derepressors. Hypoxia or anemia themselves, however, are not particularly potent in the production of hypertrophy. If this mechanism operates, it probably does so by an indirect means.

MULTIPLICATION OF SUBCELLULAR ORGANELLES AND STRUCTURES

Cardiac muscle cells are composed of a large variety of structures including the nucleus, myofibrils, mitochondria, sarcotubular systems, microsomal fraction, and subcellular enzyme systems. To add to the complexity, several cell types with different subcellular architecture are included in heart tissue. Following a stimulus for hypertrophy, all components of the cardiac muscle cell, except the nucleus, increase in size and number. However, it is still not established whether cardiac muscle cell components increase in parallel or whether synthesis of one or more structures precedes or lags behind the others. Our preliminary studies show that in cardiac hypertrophy, incorporation of radioactive amino acids is increased in all cell fractions. Some structures, however, such as mitochondria and nuclear proteins, appear to be made at more rapid rates than others in early hypertrophy. In long-term hypertrophy, it is quite conceivable that the replenishment of some of the organelles and structures may be inadequate, and cardiac decompensation may ensue. Some data have been presented to suggest that mitochondrial mass may be relatively diminished in chronic hypertrophy with congestive heart failure.[4]

Much more work is necessary for an understanding of how various cell components are synthesized and how optimal relationships between cell components are maintained. Our group has focused some of its attention on the mechanism and control of formation and destruction of one cellular organelle, the mitochondrion. The mitochondrion is the site of oxidation of substrate and of the formation of ATP, which is then used for muscle contraction and for other energy-requiring processes of the cell. Almost all energy of the cardiac muscle cell is derived from oxidative metabolism, emphasizing the vital importance of mitochondria in cardiac metabolism. About 20 per cent of cardiac muscle volume is occupied by mitochondria. In experimental cardiac hypertrophy there is a rapid proliferation of mitochondria, as indicated by a rise in mitochondrial oxidative enzymes in cytochromes. The time course of mitochondrial synthesis is now under investigation in our laboratory.

We know that organelles such as mitochondria, ribosomes, and myofibrils do not have an unlimited life, but turn over at rather rapid rates. For example, cardiac mitochondria turn over with a half-life of about 5 days,[12] as do liver ribosomes.[16] Skeletal muscle myosin has a turnover time of about 30 days.[7, 24] Data for the turnover of the latter two components in the heart are still not available. On the other hand, the cardiac muscle cell itself probably lives as long as the entire organism. Therefore even the normal cardiac muscle cell is in a state of dynamic equilibrium with regard to its components.

Organelles such as the mitochondria have a semiautonomous existence within the cell.[34, 39] The mitochondrion has its own genetic material, DNA,[22, 33] its unique species of ribosomal[36] and soluble RNA,[1] and its own system for protein synthesis.[23] Interactions between mitochondria and the cell nucleus and cytoplasm must be very complex. Such inter-

actions may be involved in the biochemical mechanism stimulating cardiac hypertrophy and must be important in the integration of the synthesis of the large variety of cell structures which form the organized cellular architecture.

POSSIBLE CAUSES OF CONGESTIVE FAILURE IN HYPERTROPHY

The previous discussion has suggested several mechanisms by which congestive failure may result from severe or long-term cardiac hypertrophy. One possibility would be that the increased work load is too great for even the most effective compensatory hypertrophy to overcome, so that decompensation results. With less severe stress it is possible that over a long period of time one or another of the various steps necessary for the synthesis of muscle components becomes limiting. Possible limiting reactions would be (1) inadequate DNA for transcription of sufficient messenger RNA; (2) inadequate RNA synthesis; (3) inadequate machinery for protein synthesis; (4) inadequate assembly or integration of organelles; and (5) inadequate synthesis of specific muscle proteins or enzymes. Deficiencies may be secondary to inadequate renewal of vital enzymes or organelles, or perhaps to the decreased efficiency related to the increase in cell size. No single primary biochemical deficiency has yet been discovered in the usual forms of congestive failure. The defect may well be on a complex organizational level. Detection may require a more complete analysis then has heretofore been made.

REFERENCES

1. Barnett, W. E., and Brown, D. H.: Mitochondrial transfer ribonucleic acids. Proc. Nat. Acad. Sci., 57:452, 1967.
2. Bartosova, D., Chvapil, M., Korecky, B., Poupa, O., Rakusan, R., Turek, Z., and Vizek, M.: The growth of the muscular and collagenous parts of the rat heart in various forms of cardiomegaly. J. Physiol., 200:285, 1969.
3. Beznak, M.: The effect of the pituitary and growth hormone on the blood pressure and on the ability of the heart to hypertrophy. J. Physiol., *116*:74, 1952.
4. Bishop, S. P., and Cole, C. R.: Ultrastructural changes in the canine myocardium with right ventricular hypertrophy and congestive heart failure. Lab. Investigation, 20:219, 1969.
5. Buccino, R. A., Harris, E., Spann, J. F., and Sonnenblick, E. H.: Response of myocardial connective tissue to development of experimental hypertrophy. Amer. J. Physiol., *216*:425, 1969.
6. Crane, W. A. J., and Dutta, L. P.: The utilization of tritiated thymidine for deoxyribonucleic acid synthesis by the lesion of experimental hypertension in rat. J. Path. Bact., *86*:83, 1963.
7. Dreyfus, J. S., Kruh, J., and Schapira, G.: Metabolism of myosin and life time of myofibrils. Biochem. J., 75:574, 1960.
8. Fanburg, B. L., and Posner, B. I.: Ribonucleic acid synthesis in experimental cardiac hypertrophy in rats. Circ. Res., 23:123, 1968.
9. Feng, T. P., and Lu, D. X.: New lights on the phenomenon of transient hypertrophy in the denervated hemidiaphragm of the rat. Sci. Sinica (Peking), *14*:1772, 1965.
10. Gilbert, W., and Müller-Hill, B.: Isolation of the *lac* repressor. Proc. Nat. Acad. Sci., *56*: 1891, 1966.
11. Goldberg, A. L.: Protein synthesis during work-induced growth of skeletal muscle. J. Cell Biol., *36*:653, 1968.

12. Gross, N. J., Getz, G. S., and Rabinowitz, M.: Apparent turnover of mitochondrial deoxyribonucleic acid and mitochondrial phospholipids in the tissues of the rat. J. Biol. Chem., *244*:1552, 1969.
13. Grove, D., Zak, R., and Nair, K. G.: Mechanisms of increase in myocardial DNA content during myocardial hypertrophy. Clin. Res., *16*:231, 1968.
14. Grove, D., Nair, K. G., and Zak, R.: Biochemical correlates of cardiac hypertrophy. III. Changes in DNA content; the relative contributions of polyploidy and mitotic activity. Circ. Res., in press.
15. Grove, D., Zak, R., Nair, K. G., and Aschenbrenner, V.: Biochemical correlates of cardiac hypertrophy. IV. Observations on the cellular organization of growth during myocardial hypertrophy. Circ. Res., in press.
16. Hirsch, C. A., and Hiatt, H. H.: Turnover of liver ribosomes in fed and fasted rats. J. Biol. Chem., *241*:5936, 1966.
17. Jacob, F., and Monod, J.: Genetic regulatory mechanisms in the synthesis of proteins. J. Mol. Biol., *3*:318, 1961.
18. Klinge, O., and Strocker, E.: Die DNS-Synthese in Rattenherzens als Function des Lebensalters. Autoradiographische Untersuchungen mit H^3-Thymidine. Experientia, *24*:167, 1968.
19. Koide, T., and Rabinowitz, M.: Biochemical correlates of cardiac hypertrophy. II. Increased rate of RNA synthesis in experimental cardiac hypertrophy in the rat. Circ. Res., *24*:9, 1969.
20. Lash, J. W., Holtzer, H., and Swift, H.: Regeneration of mature skeletal muscle. Anat. Res., *128*:679, 1957.
21. Linzbach, A. J.: Heart failure from the point of view of quantitative anatomy. Amer. J. Cardiol., *5*:370, 1960.
22. Luck, D. J. L., and Reich, E.: DNA in mitochondria of *Neurospora crassa*. Proc. Nat. Acad. Sci., *52*:931, 1964.
23. McLean, J. R., Cohn, G. L., Brandt, I. K., and Simpson, M. V.: Incorporation of labeled amino acids into the protein of muscle and liver mitochondria. J. Biol. Chem., *233*: 657, 1958.
24. McManus, I. R., and Mueller, H.: The metabolism of myosin and the meromyosin from rabbit skeletal muscle. J. Biol. Chem., *241*:5967, 1966.
25. Meerson, F. Z., Beloshapkina, T. D., Lushnikov, E. F., Leikina, E. M., Markovskaya, G. M., and Chernyshova, G. V.: Function, structure and protein metabolism of hypertrophied myocardium. Fed. Proc., *23*:T667, 1964.
26. Meerson, F. Z., and Alekhina, G. M.: The effect of contractile activity on the heart muscle cells on the DNA synthesis in the interstitial cells of myocardium. Dokl. Acad. Sci. U.S.S.R., *173*:1222, 1967.
27. Meerson, F. Z., Alekhina, G. M., Aleksandrov, P. N., and Bazardjan, A. G.: Dynamics of nucleic acid and protein synthesis of the myocardium in compensatory hyperfunction and hypertrophy of the heart. Amer. J. Cardiol., *22*:337, 1968.
28. Morkin, E., and Ashford, R. P.: Myocardial DNA synthesis in experimental cardiac hypertrophy. Amer. J. Physiol., *215*:1409, 1968.
29. Moroz, L. A.: Protein synthesis activity of heart microsomes and ribosomes during left ventricular hypertrophy in rabbits. Circ. Res., *21*:449, 1967.
30. Nair, K. G., Cutilletta, A. F., Zak, R., Koide, T., and Rabinowitz, M.: Biochemical correlates of cardiac hypertrophy. I. Experimental model; changes in heart weight, RNA content, and nuclear RNA polymerase activity. Circ. Res., *23*:451, 1968.
31. Nowy, H., Frings, H. S., and Rey, K.: On the nucleic acid content of the normal hypertrophied rabbit hearts. Experientia, *15*:70, 1959.
32. Ptashne, M.: Isolation of the phage repressor. Proc. Nat. Acad. Sci., *57*:306, 1967.
33. Rabinowitz, M., Sinclair, J., DeSalle, L., Haselkorn, R., and Swift, H.: Isolation of deoxyribonucleic acid from mitochondria of chick embryo heart and liver. Proc. Nat. Acad. Sci., *53*:1126, 1965.
34. Rabinowitz, M.: Extranuclear DNA. Bull. Soc. Chim. Biol., *50*:311, 1968.
35. Reznik, M.: Thymidine-H^3 uptake by satellite cells of regenerating skeletal muscle. J. Cell. Biol., *40*:568, 1969.
36. Rifkin, M. R., Wood, D. D., and Luck, D. J. L.: Ribosomal RNA and ribosomes from mitochondria of *Neurospora crassa*. Proc. Nat. Acad. Sci., *58*:1025, 1967.
37. Ring, P. A.: Myocardial regeneration in experimental ischemic lesions of the heart. J. Path. Bact., *62*:21, 1950.
38. Roberts, J. T., Wearn, J. T., and Boten, I.: Quantitative changes in capillary-muscle relationship in human hearts during normal growth and hypertrophy. Amer. Heart J., *21*: 617, 1941.
39. Roodyn, D. B., and Wilkie, D.: The Biogenesis of Mitochondria. London, Methuen and Company Ltd., 1968.
40. Sandritter, W., and Scomazzoni, G.: DNA content (Feulgen photometry) and dry weight (interference microscopy) of normal and hypertrophic heart muscle fibers. Nature, *202*:100, 1964.

41. Schreiber, S. S., Oratz, M., Evans, C., Silver, E., and Rothschild, M. A.: Effect of acute overload on cardiac muscle mRNA. Amer. J. Physiol., *215*:1250, 1968.
42. Shafiq, S. A., Goricki, M. A., and Mauro, A.: Mitosis during postnatal growth in skeletal and cardiac muscle of the rat. J. Anat., *103*:135, 1968.
43. Stenger, R. J., and Spiro, D.: The ultrastructure of mammalian cardiac muscle. J. Biophys. Biochem. Cytol., *9*:325, 1961.
44. Stockdale, F. E., and Holtzer, H.: DNA synthesis and myogenesis. Exper. Cell Res., *24*: 508, 1962.
45. Sumner, R. G., and McIntosh, H. D.: The nucleic acid studies in experimental cardiomegaly. Circ. Res., *12*:170, 1963.
46. Yaffe, D., and Gershon, D.: Multinucleated muscle fibers: Induction of DNA synthesis and mitosis by polyoma virus infection. Nature, *215*:421, 1967.
47. Zak, R., Grove, D., and Rabinowitz, M.: DNA synthesis in rat diaphragm as an early response to denervation. Amer. J. Physiol., *216*:647, 1969.

950 East 59th Street
Chicago, Illinois 60637

The Cellular Action of Digitalis Compounds

HANS H. HECHT, M.D.*

The action of digitalis and digitalis-like compounds on the performance characteristics of cardiac muscle—so apparent to clinicians at the bedside and to physiologists in the laboratory—still eludes all efforts to form a unified concept which should define the mechanism by which the many well studied effects can be brought into relation with each other. The obvious salutary action of digitalis in a patient with atrial fibrillation and a fast ventricular response, or the hemodynamic improvement in raising cardiac output and lowering passive pulmonary hypertension and end-diastolic ventricular pressure and volume suggest a direct effect of these compounds on cardiac contractility and on the mechanisms which control repetitive responses and heart rate. These "inotropic" and "chronotropic" effects have been well studied and have provided a broad experimental basis of digitalis action.[1, 2, 3]

The effects of digitalis compounds can be demonstrated to involve properties of cardiac muscle that are the result of complex interactions by themselves, such as contractility, excitability, alterations of the refractory period, and neural control of impulse formation and impulse conduction—not to mention the influence of such drugs on renal function and on the tone of peripheral vasculature. While these observations have advanced our knowledge of digitalis action over bedside observations and while they have pointed to excesses and the nature of intoxication, they have not delineated the precise mode of action of these compounds. Detailed studies relating chemical structure and function, while again revealing many quantitatively important factors involved in activity, speed of onset, and duration of effect,[1, 35] have also failed to yield an answer on the mechanism of action.

More recently, studies concerned with the biologic activities of the cell and cellular substructures have begun to provide some insight into the basic mechanisms of cardiac function and have added some informa-

*Professor of Medicine and Physiology, University of Chicago, Chicago, Illinois

Supported in part by grants from Eli Lilly Company, The Chicago Heart Association, the U.S. Public Health Service (Nos. HE-05673-05 and HE-05793-03), and the Myocardial Infarction Research Unit (contract No. PH-4368133).

tion of promise in answering questions on the precise mode of drug effects.[5, 6] In the case of digitalis compounds, the ionic exchange across cell membranes and their influence on excitability and conduction, the behavior of contractile mechanisms in health and disease, and the interaction between these two areas (excitation-contraction coupling) has led to a remarkable though still incomplete concept of the mechanism and sites of their action.

THE CELLULAR EVENTS OF MYOCARDIAL EXCITATION

The electrochemical events occurring with each cardiac cycle at the cell boundary and within (the "action potentials") have commonly been examined without relation to the intracellular contractile machinery, whose function, however, is largely dependent on the former. It has also been customary to view the rhythmic changes in permeabilities of the cellular surface membranes in terms of (a) a passive process triggered in the mammalian heart by a sudden and large sodium influx into the cell interior and (b) an active process of restoration which requires energy and restores the ionic transcellular gradients (Fig. 1).

The interdependence of passive and active processes is best illustrated by the fact that the activity of the carrier mechanisms or transport substances ("pumps") which are called upon to restore the resting state of the membrane is a direct function of the intracellular concentration of Na^+ and of the transmembrane chemical and electrical potential gradient.[9, 15] The activity of this mechanism, therefore, also depends on the level of the transmembrane voltage. Similarly, passive and active transports of potassium ions (in the opposite direction and on a much smaller scale) are interconnected, as is the transport of anions (Cl^- primarily). The detailed movement of other ions, notably calcium, are less well studied and will be considered later. Little is known about the physicochemical alterations which trigger the changes in ionic permeabilities and changes in membrane conductance on excitation, or about the nature of the electrical properties of and possible fixed charges on biological membranes themselves which control ionic transports in both directions. The recent observations on special organic macrocyclic compounds have begun to focus on mechanisms of active transport and on the nature of carriers and pumps.[37, 50]

The surface membrane of cardiac fibers has many infoldings and is therefore much larger than the over-all circumference of a fiber would suggest. The skin of the cell is wrinkled. Its folds (wrinkles) reach deep into the cell interior and allow the transcellular ion exchange processes discussed above to occur very close to the contractile proteins. These infoldings (transverse canaliculi) may be considered an "intracellular" network in ready communication with the outside world. The ionic concentration of fluid within these tubular structures is similar to, though because of stagnation and smallness of the tubes not exactly like, that of the extracellular surroundings. The well developed transverse tubular

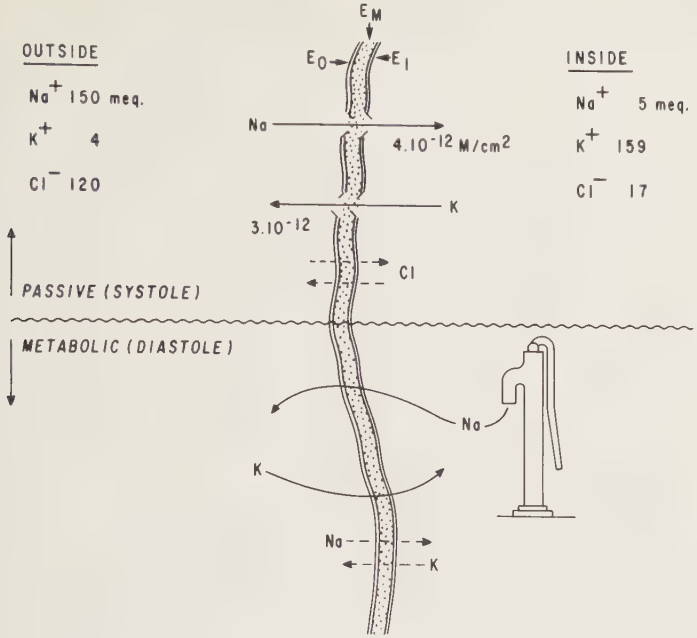


Figure 1. Simplified scheme of passive and active transport across cardiac cellular membranes.

The passive ionic transport characterizes excitation and consists of an interchange of monovalent ions along and down an "electrochemical gradient" (from high to low ionic concentration and from high to low electrical potential). The ionic movement is in the direction of restoring an electrolyte equilibrium on both sides of the membrane. This process, however, is immediately counteracted by a "pump" requiring energy (lower part of the diagram) which restores the ionic and electrical disequilibrium characteristic of the "resting" phase of the membrane.

network is the apparent reason for certain electrophysiological differences between cardiac muscle and other excitable tissues.[13]

In spite of the physical proximity of excitatory events occurring at the cell membrane to the contractile fibrillar proteins, investigators have in the past concerned themselves largely with either excitation or contraction—yet the interdependence of the two is of course essential for normal myocardial function. The coupling of the two processes is therefore of great interest. Recent studies on isolated ventricular fibers and Purkinje strands have again demonstrated the occurrence of a contraction or twitch following each action potential* or following induced changes in resting potential (depolarization).[11, 29] These studies have demonstrated a quantitative dependence of the height of the tension developed in a muscle fiber on the degree of depolarization, even if the latter is artificially induced. One will argue that the contraction is less

*The action potential is defined as a rapid change in electrical potential of the cell interior on excitation (during the rapid sodium entry) from about −80 mV. to about +20 mV., with a gradual return to the previous "resting potential" thereafter. This is the result of the electrochemical changes—active and passive—discussed above.

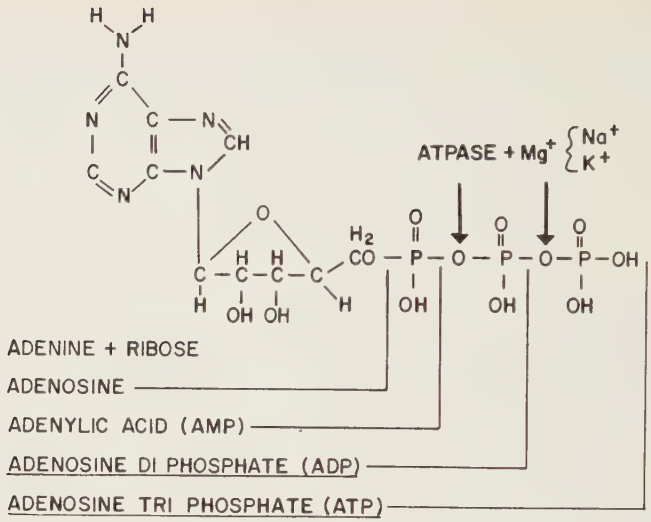


Figure 2. Energy release available for active transport processes from ATP.

Hydrolysis of a phosphodiester bond yields about 10,000 calories of free energy. The reduction of these high energy phosphate compounds requires ATPase. The membrane ATPase is active only in the presence of Mg^{++} and is blocked by digitalis glycosides.[6, 16, 48]

The major sites of ATP synthesis are mitochondria. To some extent mitochondrial metabolism is ADP-dependent, so that a certain feedback mechanism exists by which the membrane ATPase is part of a regulatory mechanism controlling mitochondrial ATP production.

than normal in examples where the Na^+ pump is more active—a debatable point.

The energy required for the active return transport mechanisms (Na^+ out, K^+ in) (Fig. 1) is largely derived from high-energy phosphate compounds, phosphocreatine and adenosine triphosphate (ATP), the latter located within the cellular membranes and largely derived from the oxidation of the citric acid cycle in the mitochondria, the subcellular "energy storehouses" particularly abundant in cardiac muscle (Fig. 2). A membrane ATPase hydrolyses the phosphodiester bonds, releasing large stores of energy and thereby reducing ATP to adenosine diphosphate (ADP) and further to adenylic acid (AMP). This enzyme is crucial for driving the mechanism of excitation and recovery and is therefore necessary in the intact organism for excitation-contraction coupling as well. The membrane ATPase functions in the presence of a potential difference across the membrane, and requires both Na^+ and K^+ for activation. The membrane ion exchange processes discussed so far depend on the presence of bivalent cations which seem to act as modifiers for the transport of monovalent cations (Na^+, K^+) (active as well as passive). In this case the membrane ATPase functions only in the presence of the bivalent cation Mg^{++}.[47, 48] The true nature of the ATPase is unknown.

THE CELLULAR EVENTS OF MYOCARDIAL CONTRACTION

Several recent reviews have dealt with the nature and function of contractile proteins, the work machinery for skeletal, smooth, and car-

diac muscle.[10, 22-24, 27] In general, at least four proteins are required to initiate and maintain contraction, two of which may be considered primary and two secondary. The primary proteins are arranged in long chains of partially overlapping filaments—a heavy protein (myosin) flanked by longer and thinner filaments in a hexagonal pattern, with myosin at the center. The thin filaments contain actin, a protein which interacts with myosin—the two constituting the primary contractile elements. The thick myosin filaments sprout connecting cross-bridges which represent the one (thick) end of the myosin molecule itself, long thin strands of the molecule forming the body of the myosin filament. Contraction occurs by sliding of the actin-containing filaments past the myosin, the consequence of an electrochemical interaction between the cross-links of myosin and the acceptor sites of thin actin filaments. The cross-bridges affect a temporary but firm union between the sets of fibrils (actomyosin complex).

It has been suggested recently that a tight attachment of the globular head of the myosin protein (the "heavy" meromyosin portion of the cross-links) to actin occurs and that sliding is effected by changes in the angle of attachment of the two adjacent proteins.[24] In any case, a link between actin and myosin is required which generates the contractile force. The energy needed for the process is provided by splitting of ATP, physically located within the myosin cross-bridge, to ADP, phosphorus ions (P_i–), and free energy. It is acted upon by an ATPase (different from the membrane ATPase discussed above) also located within the cross-bridge.[10] In turn the resulting ADP can be rephosphorylated to ATP by free cytoplasmic ATP. Shortening, tension, contraction, and work in any muscle are then the consequence of the continuous making and breaking of cross-links between the two primary contractile proteins in a ratchet-like fashion dependent on a continuous breakdown and reconstitution of ATP.

Ionized calcium is required to initiate, control, and maintain this complex interaction in vivo. It is the only physiologically available intracellular ion which causes fiber shortening when injected into the cell.[20, 40] In its absence, or in amounts of less than 10^{-7} M, the primary protein interaction does not occur and contraction is weak or absent. Since, however, highly purified actomyosin remains in the active contractile state (with high ATPase activity) in the absence of calcium,[43] a second set of proteins, "calcium sensitizers," have been demonstrated on which calcium acts directly—it does not act on the primary proteins, ATP synthesis, or the myosin ATPase. These secondary proteins, tropomyosin and troponin, are contained in the thin (actin) filaments of the system.[11, 12, 26] The mechanism appears to be that ionized calcium combines with troponin to form a tropomyosin-Ca-troponin complex at optimal pH (another factor!), which in the uncombined form covers the reactive sites of meromyosin cross-bridges and actin fibrils. The cover thus prevents the ATPase from providing the needed contractile energy. The calcium complexing with the secondary (regulatory) proteins removes this cover and opens the door for ATP splitting and interaction between the primary proteins. This mechanism has recently been reviewed and its relation to abnormal contractile states discussed and illustrated.[27, 28]

EXCITATION-CONTRACTION COUPLING[7, 30, 44]

In the heart, the intracellular concentration of ionized calcium must fluctuate with each beat: it must be high following excitation to couple mechanical work to the electrical event, and it must be removed by some "relaxing factor" to terminate the active contractile state.[51, 52] This calcium pulsing may be accomplished by periodic changes in the state of ionization and consequently changes in the ratio of ionized to bound intracellular Ca^{++} or by an actual delivery and removal of Ca^{++} to the reactive sites. The latter is strongly suggested by the demonstration of calcium storage sites in skeletal and cardiac muscle within the so-called longitudinal tubular system, the sarcotubular or endoplasmic reticulum.

In contrast to the transverse tubules, the infoldings of the outer membrane, the sarcotubular reticulum is a truly intracellular structure and forms special intimate connections (diads) at the plasma membrane facing the intracellular spaces, at the intercalated disks and at the transverse invagination of the membrane, the transverse tubular system. These extensive connections of an otherwise not too plentiful sarcotubular system are a characteristic of mammalian heart muscle.[42]

One must further consider that one of the possible factors periodically releasing calcium from the storage site is excitation (rapid sodium entry) or any mechanism which depolarizes the cell membrane. There is an interrelation between Ca^{++} and Na^{+} to the extent that passive Na^{+} entry cannot occur unless and until Ca^{++} is moved from the site of Na^{+} passage: Ca^{++} seals the resting membrane. There is a quantitative relationship as well: a high Na^{+} influx causing a large accumulation of intracellular Na^{+} is accompanied by large release of Ca^{++} and vigorous contraction. This may be a passive cation exchange mechanism within the sarcotubular system.

The details of Ca^{++} delivery from the depot to the receptor proteins are lacking and an additional relationship must be considered: the force of contraction (in heart muscle only) is strongly dependent on the ratio of $[Ca^{++}]/[Na^{+}]^2$ in the extracellular space: contraction might be increased either by the addition of Ca^{++} to the extracellular space or by the removal of Na^{+} therefrom.[6, 34, 54] For whatever reason, it is the competition between Na^{+} and Ca^{++} or the competition of these ions for a receptor site or anion radical group which causes the eventual intracellular release of ionized calcium. This initiates contraction. It is the change in Na^{+} distribution between intracellular and extracellular space coincident with excitation and depolarization which starts the event. It might in fact be postulated that the displacement of Ca^{++} from within the membrane into the sarcoplasmic reticulum is the primary change which triggers Na^{+} entry on excitation.

The removal of Ca^{++} from its site of action on the contractile proteins terminates the active contractile state: the foregoing processes are reversed, but now ionic movements are uphill and require energy—again largely derived from ATP. Ca^{++} is pumped back into the sarcoplasmic reticulum with the aid of an ATPase requiring Mg^{++},[51] and excess intra-

cellular Na^+ is pumped into the extracellular space using energy obtained from a membrane ATPase dependent on Na^+ and K^+ (Fig. 1). At least three types of ATPases are involved in these interrelated energy-requiring processes connected with (a) relaxation and (b) repolarization: myosin ATPase, sarcotubular ATPase, and membrane ATPase—all working at different sites, on different substrates, with different coenzymes, and blocked by different agents.

WHERE DOES DIGITALIS COME IN?

So far, no mention has been made of where digitalis and digitalis-like glycosides might intervene to improve myocardial contractility. The striking similarity of the action of these glycosides to the effects of calcium suggests some link between the two. O. Loewi expressed this clearly 50 years ago: "I am convinced that the effect of digitalis compounds consists of sensitizing the heart to calcium; what we assume to be the digitalis effect is in reality an increased effectiveness of calcium brought about by the action of digitalis compounds . . ."[32] Yet this connection has been elusive: a direct action of digitalis on the actomyosin complexing, the myosin ATPase, or the troponin-tropomyosin reaction is ruled out. The delivery of ATP from mitochondria is not influenced. There is no direct evidence to suggest that glycosides facilitate the ion exchange processes or the release of Ca^{++} from the sarcotubular reticulum, although these changes occur as the consequence of changes in Na^+ content; the calcium pump which restores calcium to the depot, thus initiating relaxation, is only slightly blocked by digitalis compounds.[31]

It is interesting that in fact these agents with their outstanding effect on improving impaired mechanical function have their primary site of action on sites within the cell membrane which is the seat of the major electrical changes associated with each heart beat. This is argued in a somewhat roundabout way:

The electrocardiographic effect of digitalis action, apart from irregularities and atrioventricular conduction anomalies, consists of a reduction of the QT or QU interval and a depression of the RST segment with a U-shaped deformation of T. Translated into the language of cellular electrophysiology, this implies a shortening of the action potential duration and a change in the slope of repolarization, notably of the plateau (phase 2).[21] Changes in the concentration of Na^+ or Ca^{++} in the extracellular space of the magnitude under consideration here cannot account for such changes (excessive variations in Ca^{++} cause alterations in QT and in the duration of an action potential, but without ST segment alterations or slope changes).

However, even minor changes in the concentration of extracellular potassium are followed by striking alterations in a surface electrocardiogram and in the configuration and time course of transmembrane potentials. These are not identical or even similar to the electrocardiographic changes seen with therapeutic doses of digitalis. Yet there exists a close relationship: Calhoun and Harrison[8] demonstrated in 1931

that the potassium content of cardiac muscle of dogs decreased according to the amount of digitalis administered. Potassium in dry tissue of the right and left ventricle measured 1.56 per cent and 1.44 per cent respectively in the controls (10 dogs), but 1.38 per cent and 1.40 per cent with therapeutic doses (10 dogs), and 1.07 per cent and 1.03 per cent with toxic doses of the drug (4 dogs).

Much later, Woodbury and Hecht,[55] measuring transmembrane potentials of single cardiac fibers before and after toxic doses of cardiac glycosides, demonstrated for frog ventricular muscle a marked reduction in the duration of the action potential, which they related to the clinically observed shortening of the QT interval in the surface electrocardiogram on the one hand and to similar records obtained under the influence of a high extracellular potassium content on the other (Fig. 3). Later experiments usually carried out in mammalian tissues (frog heart muscle does not have a sarcotubular system!) and in Purkinje fibers have generally demonstrated lengthening of the action potential under the influence of therapeutic doses (Fig. 4). This was related to a state of cellular and extracellular potassium depletion, shown early by Harrison and lately demonstrated forcefully as a common denominator in digitalis intoxication by Lown and his associates.[33]

The data are less confusing than it appears at first glance. Depletion of cellular potassium under the influence of digitalis therapy is suggested by the data of Calhoun and Harrison and by Lown, and others.[17, 21, 33, 38, 39, 49] Lengthening of the action potential was also noted by Woodbury and Hecht and by Kassebaum[25] during the earlier phases of the action, and a differential quantitative difference has also been reported, with ventricular fibers being more sensitive to digitalis action than the Purkinje fiber of the conduction tissue.[21] (Because of convenience of impalement of the microelectrode, the larger Purkinje tissue has been used almost exclusively in later studies.) It appears, then, that a state of cellular hypokalemia precedes a later stage of accumulation of

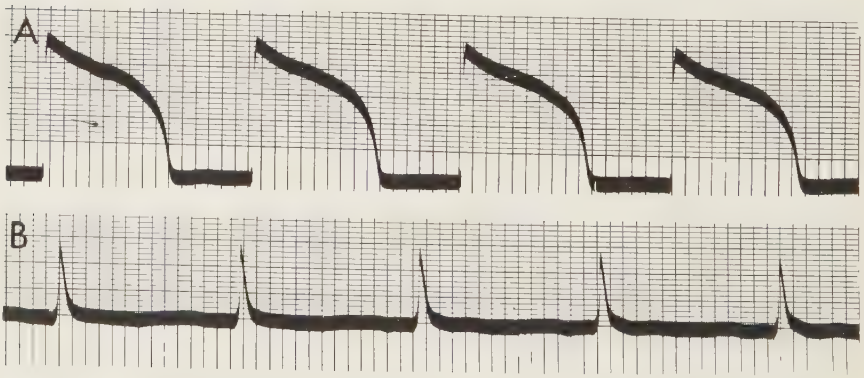


Figure 3. Intracardiac injection of a cardiac glycoside on the configuration of a cardiac action potential (frog ventricular muscle).[55]

A, Control; *B*, 15 minutes after applications of 1.2 mg. per kg. digitoxin diluted in Ringer's solution. Note marked alteration in the shape and shortening of the repolarization event (see text).

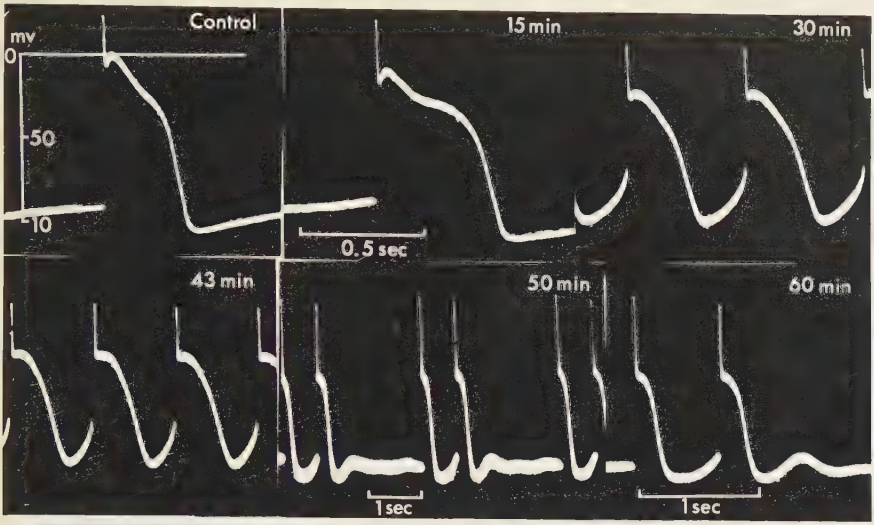


Figure 4. Effect of cardiac glycosides in therapeutic dosage on the action potentials of a sheep Purkinje fiber.

Note the early lengthening of the action potential (15 minutes) and the appearance of spontaneous rhythmicity ("ventricular tachycardia") at 30 minutes. Later oscillatory phenomena resulting in bigeminal rhythm appear (note change in film speed at 50 and 60 min.). These findings are similar to those observed in cellular potassium depletion.

Sheep, Purkinje fiber. Tyrode solution (K^+ 4.0 mEq.) with ouabain, 5×10^{-7} gm. per ml., added to the perfusate.

potassium at the outer surface of the membrane—i.e., in the extracellular space. Coincident diuresis might remove the over-all amount of extracellular K^+ present. Gross hypokalemia might or might not be present, but the innermost layer of extracellular fluid adjacent to the cellular membrane eventually accumulates K^+.

As stated above, the basic excitatory process consists of a rapid inflow of Na^+ followed by a lesser outflow of K^+ and a reversal of this process by an active metabolic pump. An accumulation of K^+ on the outer surface of the membrane (a) encourages further outflow of K^+ (enhances K^+ conduction) and (b) suggests that the mechanism responsible for re-entry of K^+ into the cell, and simultaneously for removal of intracellular Na^+, has become impaired. While the over-all K^+ content of tissues might have decreased (and this in turn is known to be followed by increased contraction), the new distribution is such that some relative excess of K^+ remains at the outer cell boundary. The paradox is thus explained. The argument would require that these digitalis-induced changes should pari passu cause an accumulation of Na^+ within the intracellular space. The classic illustration of E. Page[41] confirms this clearly (Fig. 5).

These observations and arguments lead to the obvious conclusion that digitalis compounds must act largely by interfering with the active intramembrane transport mechanism, the ion pump, very likely by blocking the membrane ATPase-dependent energy release. Quite in-

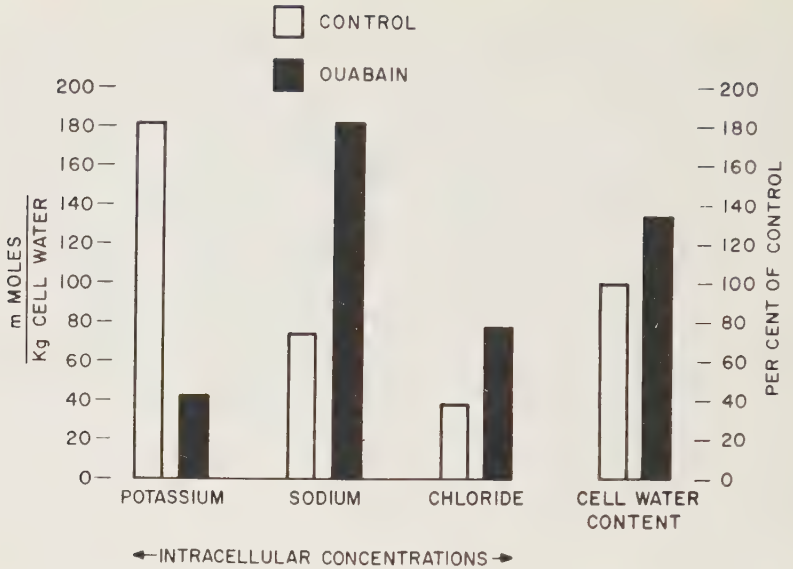


Figure 5. Effect of digitalis on intracellular water content and ionic concentrations.[41] At right ventricular papillary muscle, 3 hours in 0.01 mM ouabain.
Note accumulation of Na^+ and loss of K^+ with little if any change in the water content of the cell upon administration of digitalis compounds.

dependently this was directly demonstrated by Schatzmann in an ingenious experiment using red cell ghosts as membranes (Fig. 6).[45, 46] By arresting metabolic activity during cold storage, such cells accumulate Na^+ on the inside and K^+ on the outside of the membrane. On rewarming, the "pump" begins to function and within a few hours normal electrolyte concentration is re-established. This is effectively blocked by digitalis—one of the most direct demonstrations of a major cellular function of digitalis and digitalis-like compounds, including some unsubstituted lactones. It has been amply confirmed.[5, 16] Cardiac glycosides may inhibit sodium transport by impairing the ATPase activated by Na^+ and K^+ in other biological membranes as well: marked natriuresis and loss of urinary electrolyte concentrating capacity, coupled with decreased enzyme activity, was recently demonstrated following digoxin infusion into canine renal arteries.[36]

The sequence of events may thus be as follows: (a) redistribution of electrolytes on depolarization, (b) ion exchange with Ca^{++} release from sarcotubular reticulum, (c) Ca^{++}-tropomyosin-troponin reaction uncovering active sites for ATP-induced interaction of actin and myosin: contraction, (d) relaxation by active reabsorption of Ca^{++} into the sarcotubular storage depot, thus drawing a curtain of secondary proteins over the sites of primary protein interaction of the fibrils, (e) active removal of Na^+ and re-entry of K^+ by membrane pumps. The metabolic sodium pump is stimulated by the accumulation of intracellular Na^+; its activity intensifies as soon as depolarization occurs. It is again and further

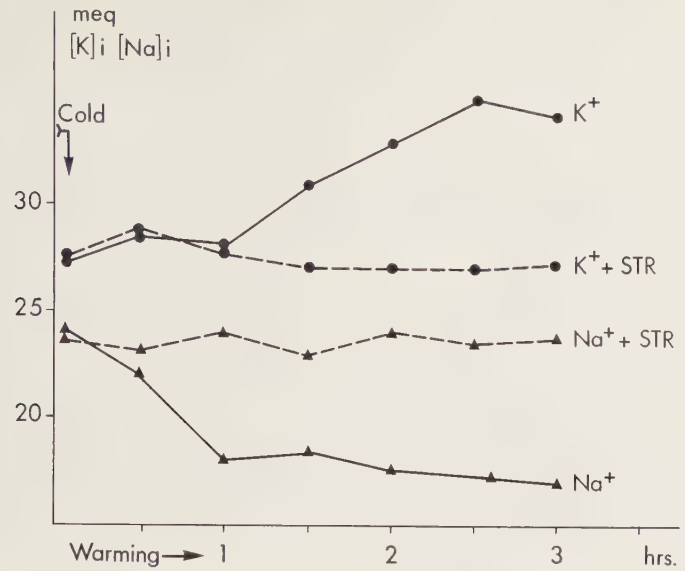


Figure 6. Effect of digitalis compounds on the active transport of ions across cell membranes (after Schatzmann[45]).

When exposed to prolonged cooling, metabolic processes are arrested, and Na^+ accumulates on the inner side of the membrane, K^+ remains outside, and intracellular K^+ concentration is low. On warming, the metabolic activity is gradually restored, and previous concentrations are reestablished. In the presence of a digitalis compound (*STR*, Strophosid), these metabolic processes are inhibited—the "pumps" are blocked, and the intracellular concentrations do not change on warming (experiments with red cell ghosts as active membranes).

stimulated when Ca^{++} begins to re-enter the storage depot, thereby releasing Na^+ into the intracellular space. Movements of other ions, notably K^+ and Cl^-, have no direct effect on contractility, but may influence it by their link to the sodium exchange.

The effect of digitalis is then concerned with (a) blocking removal of Na^+ from the intracellular space by poisoning the membrane ATPase, (b) establishing a high intracellular Na^+ concentration, making increased ions available for exchange and causing release of Ca^{++}, (c) slow return of Ca^{++} to the storage area, largely because of high Na^+ concentrations there, (d) blocking of re-entry of K^+ with consequent cellular K^+ depletion, and finally or in other regions (e) accumulation of K^+ at the outer membrane boundary with excess of K^+ conductance even in the phase of metabolic slowdown. Some of these steps are indicated in Figure 7.

While this scheme of cellular action of digitalis compounds has many attractive and experimentally verifiable features, it is still shaky as a hypothesis and leaves many questions unanswered. A physician will be concerned intuitively by the fact that such a remarkably salutary positive action, restoring contractility to normal in a malfunctioning organ, should be based primarily on its demonstrated effect on blocking a needed metabolic process. Since abnormal cardiac function and many aspects of the failing heart likewise coincide with intracellular Na^+

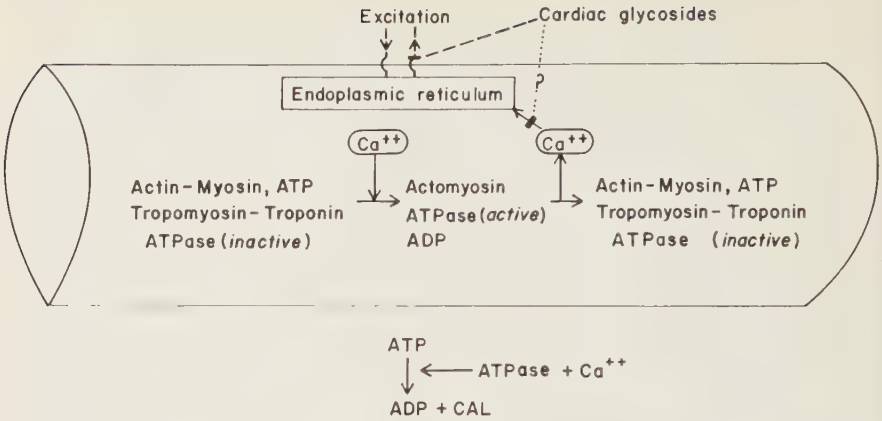


Figure 7. Scheme of excitation-contraction coupling with possible sites of digitalis action. The contractile reactions within the fiber represent a classical example of conversion of biochemical energy to mechanical work. Its relation to the electrical changes occurring with each heart beat within and at the surfaces of the fiber membrane is mediated via release and uptake of ionized calcium by the endoplasmic (or sarcoplasmic) reticulum (see text).

accumulation, differences between this and digitalis action must be delineated. This cannot be done in detail except by stating that in heart failure additional steps are involved in the process of energy conversion to contractility. It has not been determined whether alterations in cellular calcium or changes in the possible distribution of calcium pools might contribute to the mechanism of contractile failure of the intact heart. The very fact that digitalis compounds do not influence the complex scheme at all except at the membrane level speaks in favor of its action on this one site only. One cannot go further until (a) the nature of membrane pumps is further clarified—attempts in this direction were cited—(b) the membrane ATPase is further identified and its interaction with digitalis glycoside understood on a stoichiometric basis, (c) the nature of the ionic transfer reaction of the sarcotubular system is established, and (d) other intracellular reactions possibly responsible for the rhythmic release of ionic calcium are either ruled out or further specified.

The coincident effect of digitalis on the potassium membrane system presumably carries the major responsibility for its toxic effects: the changes in diastolic depolarization and direct depolarization by digitalis as well as changes in the duration of the excitatory state and the refractory period can best be interpreted as secondary to changes in K^+ distribution. Low intracellular and extracellular K^+ favors the appearance of new centers of impulse formation, ectopic foci, and re-entrant mechanisms (Fig. 4).[18, 19, 21] The appearance of atrioventricular conduction effects, independent of the action of digitalis on the acetylcholine release system, may at least in part be the consequence of accumulation of K^+ on the outer membrane surface.

We seem to know much but not enough; we seem to understand some of the parts, but cannot fully comprehend the entire clue of the witches brew entrusted to Whithering by the old lady 200 years ago.[53]

"Hab ich die Teile in meiner Hand, fehlt leider nur das geistige Band." ("I have the parts and know them well—but the whole of the story I cannot tell.") (Goethe, Faust I).

REFERENCES

RECENT REVIEWS

1. Doherty, J. E.: Clinical pharmacology of the digitalis glycosides: A review. Amer. J. Med. Sci., *255*:382, 1966.
2. Koch-Weser, J.: Mechanism of digitalis action on the heart. New Eng. J. Med., *277*:416, 469, 1967.
3. Luchi, R. J., and Conn, H. L.: Digitalis action on the cells: Fable, fact and fancy. Prog. Cardiovasc. Dis., *7*:336, 1965.
4. Mason, D. T., Spann, J. F., and Zelis, R.: New developments in the understanding of the action of digitalis glycosides. Prog. Cardiovasc. Dis., *11*:443, 1969.
5. Repke, K.: Über den biochemischen Wirkungsmodus von Digitalis. Klin. Wochenschr., *42*:157, 1964.
6. Wilbrandt, W.: Mechanism of action of cardiac glycosides. *In* Briller, S. T., and Conn, H. L., eds.: The Myocardial Cell. Philadelphia, University of Pennsylvania Press, 1966, p. 297.

GENERAL REFERENCES

7. Brady, A. J.: Excitation and excitation-contraction coupling in cardiac muscle. Ann. Rev. Physiol., *26*:341, 1964.
8. Calhoun, J. A., and Harrison, T. R.: Studies in congestive heart failure. IX. The effect of digitalis on the potassium content of the cardiac muscle of dogs. J. Clin. Invest., *10*:139, 1931.
9. Conway, E. J.: Critical energy barriers in the excretion of sodium. Nature, *187*:394, 1960.
10. Davies, R. E.: Molecular theory of muscle contraction: Calcium-dependent contractions with hydrogen bond formation plus ATP-dependent extensions of part of the myosin-actin cross-bridges. Nature, *199*:1068, 1963.
11. Ebashi, E., Ebashi, F., and Kodoma, A.: Troponin. I. Preparation and physiological function. J. Biochem., *62*:137, 1967.
12. Ebashi, S., and Kodoma, A.: Native tropomyosin-like action of troponin on trypsin treated myosin B. J. Biochem., *60*:733, 1966.
13. Fozzard, H. A.: Membrane capacitance of the cardiac Purkinje fibre. J. Physiol., *182*:255, 1966.
14. Fozzard, H. A., and Hellam, D. C.: Relationship between membrane voltage and tension in voltage-clamped cardiac Purkinje fibers. Nature, *218*:588, 1968.
15. Fozzard, H. A., and Kipnis, D. M.: Regulation of intracellular sodium concentrations in rat diaphragm muscle. Science, *156*:1257, 1967.
16. Glynn, I. M.: The action of cardiac glycosides on sodium and potassium movements in human red cells. J. Physiol., *136*:148, 1957.
17. Hajdu, S., and Leonard, E.: Cellular basis of cardiac glycoside action. Pharmacol. Rev., *11*:173, 1959.
18. Hecht, H. H.: Electrical properties of cardiac fibers. *In* Cyclopedia of Medicine, Surgery and Specialties. Philadelphia, F. A. Davis Co., 1964, Part I, Vol. III, pp. 55-73.
19. Hecht, H. H.: Electrophysiologic principles of cardiac fibers. *In* Gould, S. E., ed.: Pathology of the Heart and Blood Vessels. Springfield, Illinois, Charles C Thomas, 1968, Chapter V.
20. Heilbrunn, L. V., and Wiercinski, F. J.: Action of various cations on muscle protoplasm. J. Cell Comp. Physiol., *29*:15, 1947.
21. Hoffman, B. F., and Singer, D. H.: Effects of digitalis on electrical activity of cardiac fibers. Prog. Cardiovasc. Dis., *7*:226, 1964.
22. Huxley, A. F., and Niedergerke, R.: Structural changes in muscle contraction: Interference microscopy of living muscle fibers. Nature, *173*:971, 1954.
23. Huxley, H. E.: Contraction of muscle. Scientific American, November 1958, p. 67.
24. Huxley, H. E.: Mechanism of muscular contraction. Science, *164*:1356, 1969.
25. Kassebaum, D.: Electrophysiological effects of strophantin in the heart. J. Pharmacol., *140*:329, 1963.

26. Katz, A.: Purification and properties of a tropomyosin-containing protein fraction that sensitizes reconstituted actomyosin to calcium-binding agents. J. Biol. Chem., *421*: 1522, 1966.
27. Katz, A.: Contractile proteins of the heart. Physiol. Rev., in press.
28. Katz, A. M., and Hecht, H. H.: Early pump failure of the ischemic heart. Amer. J. Med., *47*:497, 1969.
29. Kavaler, F.: Membrane depolarization as a cause of tension development in mammalian ventricular muscle. Amer. J. Physiol., *197*:968, 1959.
30. Langer, G. A.: Ion fluxes in cardiac excitation and contraction and their relation to myocardial contractility. Physiol. Rev., *48*:708, 1968.
31. Lee, K. S., Yu, D. H., and Struthers, J. J.: A study on the effect of cardiac glycosides on the syneresis of myofibrils in the presence of relaxing factor. J. Pharm. Exper. Therap., *148*:277, 1965.
32. Loewi, O.: Über den Zusammenhang zwischen Digitalis- und Kalziumwirkung. Arch. Exp. Path. Pharmacol., *82*:131, 1918.
33. Lown, B., and Levine, S. A.: Current concepts in digitalis therapy. New Engl. J. Med., *250*:771, 1954.
34. Lüttgau, H. C., and Niedergerke, R.: Antagonism between Ca and Na ions on the frog's heart. J. Physiol., *143*:486, 1958.
35. Luchi, R. J.: Some biochemical reactions influenced by the digitalis glycosides. *In* Briller, S. A., and Conn, H. L., eds.: The Myocardial Cell. Philadelphia, University of Pennsylvania Press, 1966, pp. 309-330.
36. Martinez-Maldonado, M., Allen, J. C., Sknoyan, G., Suki, W., and Schwartz, A.: Renal concentrating mechanism: Possible role for sodium-potassium activated adenosine triphosphatase. Science, *165*:807, 1969.
37. Mueller, P., and Rudin, D. O.: Resting and action potentials in experimental bimolecular lipid membranes. J. Theoret. Biol., *18*:222, 1968.
38. Müller, P.: Das Elektrokardiogramm bei Elektrolytstörungen. Schweiz. Med. Wochenschr., *92*:842, 1962.
39. Müller, P.: Kalium und Digitalis Toxizität. Cardiologia, *42*:176, 1963.
40. Niedergerke, R.: Local muscle shortening by intracellular application of calcium. J. Physiol., *128*:12, 1955.
41. Page, E.: Actions of cardiac glycosides on heart muscle cells. Circulation, *30*:237, 1964.
42. Page, E.: Tubular systems in Purkinje cells of the cat heart. J. Ultrastruct. Res., *17*:72, 1966.
43. Perry, S. V., and Grey, T. C.: Ethylene diamine tetra-acetate and the adenosinetriphosphatase activity of actomyosin systems. Biochem. J., *64*:5, 1956.
44. Sandow, A.: Excitation-contraction coupling in skeletal muscle. Pharmacol. Rev., *17*:265, 1965.
45. Schatzmann, H. J.: Herzglykoside als Hemmstoffe für den aktiven Kalium- und Natriumtransport durch die Erythrocytenmembran. Helv. Physiol. Acta, *11*:346, 1953.
46. Schatzmann, H. J.: Effect of cardiac glycosides on active Na-K transport. Protoplasma, *63*:136, 1967.
47. Skou, J. C.: Enzymatic basis for active transport of Na^+ and K^+ across cell membranes. Physiol. Rev., *45*:596, 1965.
48. Skou, J. C.: Enzymatic basis for the active transport of sodium and potassium. Protoplasma, *63*:303, 1967.
49. Sulser, F.: Die Wirkung von Herzglykosiden und Kalzium auf den Kaliumtransport im Herzmuskel. Experientia, *15*:97, 1959.
50. Tostenson, D. C.: Effect of macrocyclic compounds on the ionic permeability of artificial and natural membranes. Fed. Proc., *27*:1269, 1968.
51. Weber, A., Herz, R., and Reiss, I.: On the mechanism of the relaxing effect of fragmented sarcoplasmic reticulum. J. Gen. Physiol., *46*:679, 1963.
52. Weber, A., Herz, R., and Reiss, I.: Regulation of myofibrillar activity by calcium. Proc. Royal Soc. B., *160*:489, 1964.
53. Withering, W.: An account of the foxglove and some of its medical uses with practical remarks on dropsy and other diseases. Birmingham, M. Swinney, 1785.
54. Wilbrandt, W., and Koller, H.: Die Calciumwirkung am Froscherzen als Funktion des Ionengleichgewichts zwischen Zellmembran und Umgebung. Helv. Physiol. Acta, *6*:208, 1948.
55. Woodbury, L. A., and Hecht, H. H.: Effects of cardiac glycosides upon the electrical activity of single ventricular fibers of the frog heart and their relation to the digitalis effect of the electrocardiogram. Circulation, *6*:172, 1952.

950 East 59th Street
Chicago, Illinois 60637

The Hemodynamic Effects of Myocardial Infarction and Results of Therapy

ROLF M. GUNNAR, M.S., M.D.*

HENRY S. LOEB, M.D.**

RAYMOND J. PIETRAS, M.D.***

JOHN R. TOBIN, JR., M.S., M.D.****

Until recently the hemodynamic effects of acute myocardial infarction have been inferred from animal experimentation. Although such experiments have led to a better understanding of the heart's response to acute ischemia and infarction, they cannot be related directly to the syndrome in man because of the inability to develop a model simulating diffuse coronary artery disease on which to superimpose acute myocardial infarction. It has been apparent since the experiments of Wiggers[38] that the myocardium, when ischemic, loses its ability to contract, and the ischemic area bulges during ventricular systole. If ischemia is produced by ligation of a major coronary vessel, then either increased work of viable myocardium compensates for the weakened

---

From the Department of Adult Cardiology and the Hektoen Institute for Medical Research of the Cook County Hospital; the Section of Cardiology, Department of Medicine, University of Illinois College of Medicine, Chicago; and the Department of Medicine, Loyola University Stritch School of Medicine, Hines, Illinois

*Professor of Medicine and Chief, Section of Cardiology, Department of Medicine, University of Illinois College of Medicine; Associate Director, Department of Adult Cardiology, Cook County Hospital

**Assistant Professor of Medicine, University of Illinois College of Medicine; Director, Intensive Care Unit, Cook County Hospital

***Assistant Professor of Medicine, University of Illinois College of Medicine; Director, Cardiac Catheterization Laboratory, University of Illinois Research and Educational Hospitals

****Professor of Medicine, Loyola University Stritch School of Medicine; Director, Department of Adult Cardiology, Cook County Hospital

bulging area or ventricular fibrillation occurs.[39] Until the use of propranolol this limitation of fibrillation made it impossible to produce the shock syndrome merely by ligating a coronary artery.

In order to produce the shock syndrome of myocardial infarction in the dog, Agress developed a model employing the injection of beads into the coronary arteries.[1] This model has been modified by using selective injection of the beads and mercury injection,[23] but remains the basic model for study of shock from myocardial infarction in the experimental animal. Agress found that in addition to a decreased cardiac output from myocardial infarction there was, in his model, a lack of compensatory vasoconstriction, and this led to hypotension and shock. He attributed this lack of the expected peripheral response to a reflex arising in the coronary arteries, but subsequent investigators have suggested that activation of stretch receptors in the bulging myocardium causes vasodilatation[8] and thus inhibits the compensatory vasoconstriction.

The first studies of cardiac output and intra-arterial pressure in patients with acute myocardial infarction were done by Gilbert et al.[11] The studies reported since that time have been reviewed recently by Kuhn[19] and show that the primary defect in acute myocardial infarction is a decrease in cardiac output. If this decrease in cardiac output is severe or not compensated for by a rise in vascular resistance, the shock syndrome will ensue. Hemodynamic deterioration may also appear with arrhythmias, particularly ventricular tachycardia and bradyarrhythmias of all types.[10] The information from these studies has been supplemented by studies of patients with cardiogenic shock of other types, particularly that seen following cardiac surgery, but these latter studies should not be interpreted as simulating the shock of myocardial infarction, because the coronary arteries of these patients are not necessarily diseased and their uniform peripheral vascular response is severe vasoconstriction.[22]

Cardiac output in the studies of myocardial infarction in the human has been measured by the indicator dilution technique, and in most of the recent series indocyanine green has been used as the indicator. Some of the objections to the Stewart-Hamilton method of analysis of the curve can be overcome by more central injecting and sampling sites,[18, 31] and even though very low cardiac outputs may be underestimated, changes during therapy are easily identified.

Left ventricular hemodynamic events have only recently been measured in patients with acute myocardial infarction. We have measured left ventricular pressures using a ventriculography catheter and have encountered less myocardial irritability than with right ventricular catheterization.[24] Cohn et al. have used a Keifa catheter with a "pigtail" loop for blind entry of the left ventricle and in addition to shock patients have studied patients with relatively uncomplicated infarction.[7] Rapaport has studied several patients with both complicated and uncomplicated myocardial infarction by floating a small catheter into the pulmonary artery. By assuming the end-diastolic pulmonary artery pressure to be equal to left ventricular end-diastolic pressure, he has thus been able to assess left ventricular function.[32] Shapiro et al. have measured arterial oxygen tension as a reflection of the severity of cardiac damage.[33]

The usefulness of this latter measurement in assessing cardiac damage has been supported by animal experiments indicating that the hypoxemia of myocardial infarction is due to pulmonary edema.[9] All these studies, including our own, demonstrate that covert left ventricular failure is more common than previously suspected in uncomplicated myocardial infarction.

SHOCK IN MYOCARDIAL INFARCTION

Hemodynamic measurements in acute myocardial infarction are practical and useful in directing the therapy of patients with severe circulatory insufficiency or shock. However, before one can discuss therapy of shock in acute myocardial infarction, definition must be agreed upon. Circulatory insufficiency is a clinical syndrome which includes obtundation, oliguria, and cold, cyanotic extremities. The arterial pressure is usually low when determined by the cuff method. To fulfill our criteria of shock, the intra-arterial pressure must also be reduced to a mean level below 80 mm. Hg.

If this definition is accepted, then we must recognize two other syndromes. Hypotension unassociated with clinical evidence of shock is a syndrome in which the patient remains alert, urine output is adequate, the extremities are warm, the pulse volume is good, and recovery is likely. This syndrome is not a cause for treatment, except for the use of atropine if the pulse rate is also slow. The second syndrome is circulatory insufficiency or "clinical shock" with a low arterial pressure as measured by the cuff method but with a normal or even high intra-arterial pressure. This syndrome is associated with poor tissue perfusion and a high mortality, but differs from the complete syndrome in that immediate support of arterial pressure is unnecessary. Shock, therefore, is associated with a decrease in cardiac output more profound than that seen in patients with acute myocardial infarction who are not in shock.[16] It is associated with either a normal or high systemic vascular resistance.[13] At least half of the patients presenting in shock with acute myocardial infarction have a normal systemic vascular resistance, which is an inappropriate response to the drop in cardiac output and is similar to the peripheral vascular response of Agress' dogs.[13, 19]

As the arterial pressure falls in patients with acute myocardial infarction, perfusion to the ischemic area decreases, and this may cause extension of the infarct. Since patients with acute myocardial infarction almost always have severe coronary disease involving the vessels supplying the noninfarcted myocardium,[5] there will be areas in which the coronary blood flow is pressure-dependent.[30] Thus, unlike the dog with normal coronary arteries, as arterial pressure falls in a man with acute myocardial infarction, he may not be able to benefit from the decreased pressure work because of an inordinate increase in ischemia to the ventricle as coronary perfusing pressure falls. Since anoxemia is a powerful stimulus to systemic vasodilatation,[3] the collateral vessels surrounding the area of infarction must be maximally vasodilated, and

because of this, flow in these areas also will be pressure-dependent even if the major vessel is patent. The marked increase in ischemia thus produced by hypotension is a reasonable explanation for the rapid clinical deterioration and cardiac irritability or bradyarrhythmia which we have observed in patients in shock with myocardial infarction when effective pressor agents are withdrawn.[15]

Although not usually considered to be important in the etiology of shock in acute myocardial infarction, hypovolemia may at times play a major role. Hypovolemia may even appear at the onset of the illness, particularly after an episode of circulatory arrest. The sudden severe acidosis associated with circulatory arrest may cause the loss of venomotor tone and loss of fluid from the intravascular compartment, creating an acute relative or absolute intravascular fluid deficit. Later in the illness hypovolemia may occur because of fluid loss induced by diuretics given at the onset of the illness or because of sweating or vomiting associated with pain or medication. If patients are kept on vasopressor agents for a long period, they may develop hypovolemia and shock similar to the norepinephrine-shock syndrome.[12] Longsjoen has pointed out that patients in a coronary care unit are frequently on a restricted fluid intake and over several days in the unit can develop significant negative water balance.[20] Finally, the syndrome of pressor dependence seen in the shock patient after prolonged pharmacologic support has been found to be associated frequently with hypovolemia and to be relieved by plasma volume expansion as the pressor agent is withdrawn.[17]

A final mechanism of shock in patients with acute myocardial infarction is that associated with a coincident illness. The elderly patient with myocardial infarction may need urinary bladder catheterization, and when on the second or third day of illness he develops the shock syndrome, it may be due to gram negative bacteremia rather than myocardial infarction. Since urinary tract instrumentation in the elderly is associated with a high incidence of bacteremia,[4] this sort of precipitation of shock in acute myocardial infarction is probably more common than has been recognized to this point.

PHYSIOLOGIC MEASUREMENTS AND THERAPY

The clinician faced with a patient who has acute myocardial infarction must decide which hemodynamic measurements are pertinent and what he should do with the information so obtained.

It is our opinion that all patients with acute myocardial infarction should have central venous pressures monitored during the first few days as they are being monitored for arrhythmias. An open venous line is imperative to quick treatment of arrhythmias and if a catheter is to be introduced, it can be inserted as far as the superior vena cava without increased complications. If the patient develops arterial hypotension associated with a fall in central venous pressure and bradycardia, he should be treated with atropine first, since this hypotension can represent a vagal effect. If atropine does not promptly correct the hypotension, or

if the patient has tachycardia, plasma volume expansion should be considered. With the central venous pressure below 10 cm. H_2O in the patient who has evidence of circulatory insufficiency, such as cold extremities, decreased urine output, and obtundation, one is justified in cautiously giving small increments of fluid, either dextrose or low molecular weight dextran.[2, 26, 37] The fluid should be given in 100 ml. increments over a 10 minute period and the response of the venous pressure monitored over the next 10 minutes. If there is no significant rise (greater than 1 cm. H_2O) in venous pressure, the procedure can be repeated.

We have treated 12 patients in shock associated with acute myocardial infarction using low molecular weight dextran infusion for plasma volume expansion.[26] Two patients had central venous pressures above 10 cm. H_2O and neither responded well to plasma volume expansion. The remaining 10 patients had central venous pressures below 10 cm. H_2O, and five responded favorably to dextran infusion, increased their cardiac outputs and arterial pressures, and survived. Such a salvage rate warrants the risks associated with volume expansion in acute myocardial infarction. These risks are minimized by selecting for volume expansion only patients manifesting circulatory insufficiency and a central venous pressure below 10 cm. H_2O.

Although, as a static measurement, central venous pressure may poorly reflect left ventricular end-diastolic pressure in acute myocardial infarction, changes in central venous pressure during volume expansion can be expected to reflect changes in left ventricular end-diastolic pressure.[24] In a series of 23 patients with various forms of shock and including three with myocardial infarction, changes in central venous pressure accurately paralleled changes in left ventricular end-diastolic pressure during volume expansion. This relationship, however, did not pertain if pressor agents were added or the rate of infusion was changed during volume expansion. Even with these precautions—starting with central venous pressure below 10 cm. H_2O; giving increments of fluid and stopping if the central venous pressure rises more than 1 cm. H_2O with a 100 ml. infusion; and not adjusting or adding pressor agents simultaneously—it can be anticipated that, because of the nature of the illness, pulmonary edema will develop in some patients. This is a small but acceptable risk when compared to the salvage rate for those who respond to this therapy. It can be treated promptly by fluid withdrawal through the same central venous pressure catheter.

If the patient presents with hypotension and clinical shock, the arterial pressure must be supported. In our experience norepinephrine is still the agent of choice. If the pressure immediately returns toward normal, the norepinephrine infusion can be slowed or discontinued. The arterial pressure should not be brought above 100 mm. Hg systolic. If the patient remains with evidence of circulatory insufficiency despite these efforts, then it is our opinion that intra-arterial pressure must be measured. If the arterial pulse pressure is small, there may be a large discrepancy between the cuff pressure and the actual intra-arterial pressure.[14] Cohn has reported discrepancies of as much as 164 mm. Hg.[6]

This discrepancy can be particularly great in patients receiving vasoconstrictors, and on several occasions we have seen patients become hypertensive during norepinephrine infusion and at a time when the cuff pressure was either unobtainable or recorded as very low. It is for this reason that comparisons of agents used to treat patients in shock are invalid without direct measurement of the intravascular pressure changes. The role of pressor agents in the shock of acute myocardial infarction is to establish an intra-arterial pressure sufficient to perfuse the myocardium. Since patients with this syndrome have severe arterial disease, the pressure needed will be above the levels needed to perfuse coronary arteries in patients without vascular disease.

Purely pressor agents such as methoxamine, Neo-Synephrine and angiotensin have no place in the treatment of patients with acute myocardial infarction. Even though some of these patients have an inappropriate peripheral vascular response to the drop in cardiac output, merely increasing the systemic vascular resistance causes clinical deterioration and further decrease in cardiac output and, in many cases, will not increase the arterial pressure to above shock levels.[13] Norepinephrine, on the other hand, increases cardiac output and arterial pressure when given to the patient who is in shock.[13, 15, 16] It should be given in amounts just sufficient to increase the intra-arterial pressure to a mean of 80 mm. Hg. If the obtundation clears and urine flow returns, these levels of pressure should be maintained. It is a rare instance in which the patient has been chronically severely hypertensive and pressures have to be raised above this level. We have demonstrated a decline in cardiac output as norepinephrine infusion rate is further increased in patients who had shown an increase in cardiac output, as the norepinephrine infusion raised arterial pressure from shock to near normal levels.[13] We have evaluated norepinephrine in 33 patients during shock associated with acute myocardial infarction, and in 25 of these it was possible to obtain adequate control values (Fig. 1). With the infusion of this agent, mean arterial pressure could be increased 43 per cent to an average value of 83 mm. Hg. during infusion. It can be seen from Figure 1 that in several patients the arterial pressure was raised above optimal levels, and we would consider these inappropriate levels a reflection of our early patient studies. Cardiac output increased an average of 18 per cent to a level of 3.3 liters per minute, while central venous pressure changed little, averaging an 0.7 mm. Hg increase to 12.8 mm. Hg during the infusion. The systemic vascular resistance increased 37 per cent to a level of 25.7 mm. Hg per liter per minute during the infusion.

Because of the effectiveness of isoproterenol when used to treat patients with cardiogenic shock following cardiac surgery, we have given this agent to 16 patients in shock associated with acute myocardial infarction (Fig. 2). In 12 of these patients, adequate controls could be obtained, and in 12 patients the results could be compared to norepinephrine infusion. In only one patient could the arterial pressure be elevated above shock levels; the average change in arterial pressure was +2.5 per cent to a mean level of 63 mm. Hg during the infusion. Only two patients showed evidence of clinical improvement, and one of these

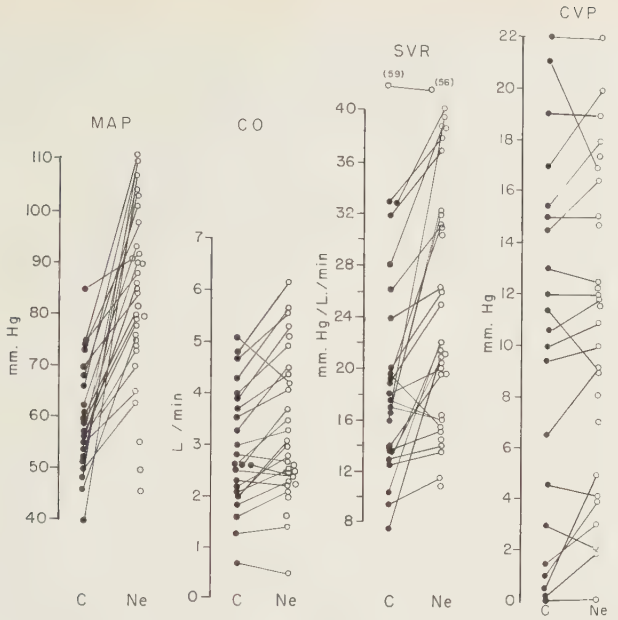


Figure 1. Results of norepinephrine infusion in patients with shock associated with acute myocardial infarction. Control measurements (C) of mean arterial pressure (MAP), cardiac output (CO), systemic vascular resistance (SVR), and central venous pressure (CVP) are shown in solid circles, (●), and measurements during infusion (NE) are shown in open circles. Lines connect measurements in individual patients. Open circles not connected by lines are measurements during drug infusion in patients in whom control values could not be obtained.

was a patient who had complete heart block; he demonstrated a further improvement with pacemaker insertion. Cardiac output during the isoproterenol infusion increased 38 per cent to a mean level of 3.6 liters per minute. Systemic vascular resistance fell 24 per cent to a mean level of 16.6 mm. Hg per liter per minute during the drug infusion.

Several patients showed rapid clinical deterioration following the switch from norepinephrine to isoproterenol infusion. We have concluded that isoproterenol is unsatisfactory in the treatment of the patient who is in shock with profound hypotension due to acute myocardial infarction.[15] It has, in our experience, been effective in some patients who had complete heart block and shock prior to inserting a pacemaker catheter. Morse et al. have shown isoproterenol to be a useful agent when accompanied by simultaneous plasma volume expansion in treating myocardial infarction patients who are mildly hypotensive and exhibit clinical shock.[29] Small amounts of norepinephrine might have been as effective in these patients, but the comparison was not made. Smith et al. reported a better response to isoproterenol than to metaraminol in patients in shock associated with acute myocardial infarction, but again, the patients who were markedly hypotensive did not show a salutary response to isoproterenol infusion.[34] Therefore, it seems that if iso-

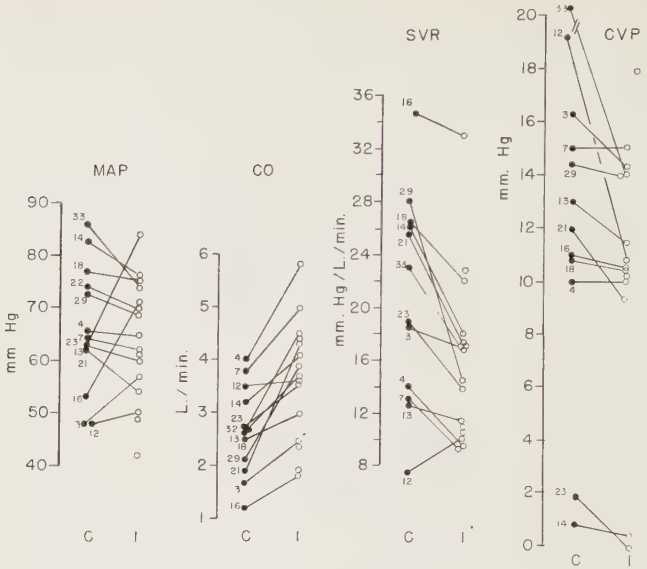


Figure 2. Results of isoproterenol infusion in patients with shock associated with acute myocardial infarction. Small numbers next to control values are to identify individual patients. Normal MAP levels in patients 33 and 14 reflect stabilization of pressure during the control period just prior to isoproterenol infusion and not initial levels at entry into the study.

proterenol has any place in the treatment of myocardial infarction, that place is in the treatment of patients who exhibit the shock syndrome but maintain normal intra-arterial pressure, and under these circumstances the infusion should be accompanied by simultaneous plasma volume expansion. It also has a place in the short period between recognition of complete heart block and shock associated with acute myocardial infarction and the institution of electrical pacing.

We have given dopamine to four patients with acute myocardial infarction (Fig. 3). In one patient the infusion had to be discontinued because of arrhythmias. One patient showed an increase in cardiac output. The mean arterial pressure did not change. In one patient urine flow doubled, but arterial pressure did not change, nor did cardiac output. The fourth patient showed decreases of both mean arterial pressure and cardiac output. This is an interesting agent, again not the agent of choice for the severely hypotensive patients, but appearing to have an intermediate action in supplying a mild amount of arterial constriction in relationship to the inotropic effect.[36]

Digitalis should be given to patients after the mean arterial pressure has been stabilized and if there is evidence of failure or elevated central venous pressure. We have evaluated digitalis in 12 patients (Fig. 4), eight of whom were receiving a constant infusion of norepinephrine during the digitalis administration. Changes in mean arterial pressure varied between −34 and +31 per cent of the control levels. Central venous pressure fell in all but four patients. In only one patient, however,

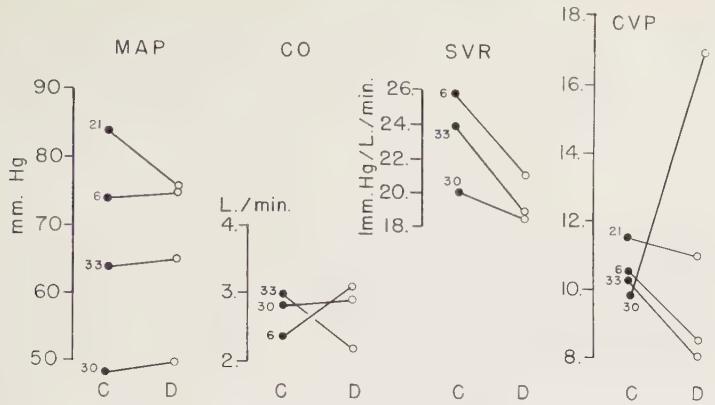


Figure 3. Results of dopamine infusion in patients with shock associated with acute myocardial infarction. Patient 21 developed ventricular tachycardia and the drug was discontinued before cardiac output could be measured.

did cardiac output increase by more than 15 per cent, and the average change from the control levels was −6 per cent.

Hemodynamic measurements during ventricular pacing for complete heart block in acute myocardial infarction have been reported,[21] and we have made measurements in three patients in shock associated with acute myocardial infarction (Fig. 5). Two patients had inferior wall infarcts, while one (patient 6) had an extensive anterior infarct. Arterial pressure increased as did cardiac output in all three patients. Patient 2 developed interference with the pacemaker after arterial pressure was stabilized at 90 mm. Hg with cardiac output of 3.9 liters per minute and

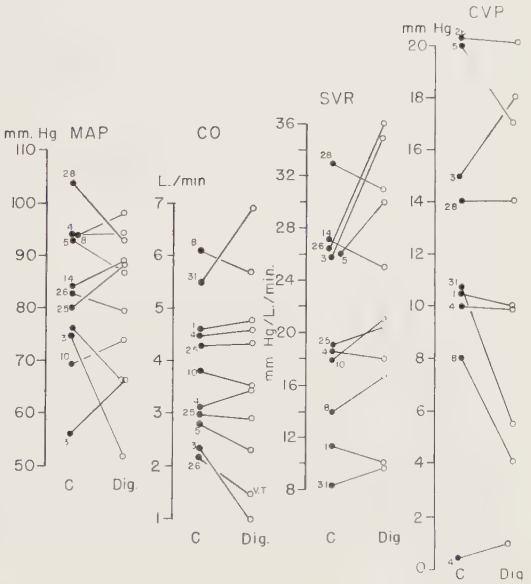


Figure 4. Hemodynamic measurements before and after digitalization in patients in shock with acute myocardial infarction. Pressor infusions were held constant during the 20 to 60 minutes between measurements.

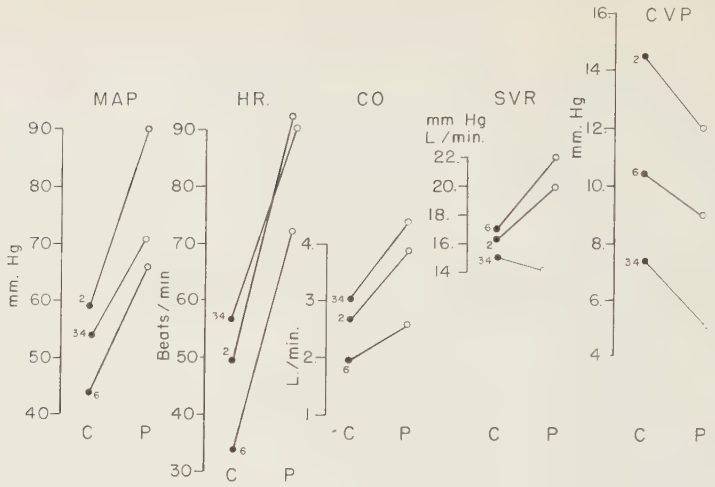


Figure 5. Results of right ventricular catheter pacing in patients in shock with acute myocardial infarction.

at a pacing rate of 94 per minute. It was then noted on several occasions that with sinus activation of the ventricle the arterial pressure increased further, and with discontinuance of artificial pacing arterial pressure stabilized at 104 mm. Hg and cardiac output rose to 4.6 liters per minute despite a fall in heart rate to 84 per minute. This is consistent with the beneficial effect of atrial contraction in supporting ventricular function in patients with damaged ventricles[28] and suggests that further trial of sequential A-V pacing is needed.[21]

Arterial hypoxemia as manifested by a decrease in the Po_2 is now being recognized more readily in patients with acute myocardial infarction. Smith et al.[35] have shown that these patients have a decreased Po_2, which is usually correctable by the administration of 100 per cent oxygen. However, in patients with complications of either severe failure or shock, 100 per cent oxygen may be insufficient unless accompanied by respiratory assistance. MacKenzie has pointed out the need for the early assessment of arterial oxygen saturation and correction of hypoxemia in the treatment of patients in shock.[27] Even though in the normal individual, the work of the heart increases when oxygen is given, it appears to be a desirable therapeutic aid in all patients with acute myocardial infarction unless one can prove that the arterial oxygen saturation is normal while the patient is breathing room air.

Vasodilators, in general, have no place in the treatment of patients with acute myocardial infarction. However, in rare instances, after the arterial pressure is stabilized, if the patient remains oliguric with cold extremities and the central venous pressure has been elevated following digitalization, a vasodilator may be tried. We have used chlorpromazine in small amounts, 5 mg. intravenously, and found this to be an effective vasodilator if the vasodilatation is accompanied by further volume expansion.[25]

Finally, for the patient with shock following massive myocardial infarction, mechanical circulatory assistance may improve chances for survival. These patients need to be selected early, and hemodynamic measurements should play a major role in such selection.

SUMMARY

In patients with acute myocardial infarction and circulatory insufficiency, hemodynamic measurements, including central venous pressure, intra-arterial pressure, and some assessment of cardiac output, are now emerging as practical and necessary for proper monitoring of therapy. More difficult measurements such as intracavitary pressures remain experimental tools, but it is hoped that the information derived will simplify the interpretation of easily obtained data, including measurements of systolic time intervals and apexcardiograms. The goal is to select and monitor pharmacologic therapy and set criteria for mechanical circulatory assistance.

REFERENCES

1. Agress, C. M., Rosenberg, M. J., Jacobs, H. I., Binder, M. J., Schneiderman, A., and Clark, W. G.: Protracted shock in the closed chest dog following coronary embolization with graded microspheres. Amer. J. Physiol., *170*:536, 1952.
2. Allen, H. N., Danzig, R., and Swan, H. J. C.: Incidence and significance of relative hypovolemia as a cause of shock associated with acute myocardial infarction. Circulation, *36*, Suppl. II, 1967, p. 50.
3. Berne, R. M.: Regulation of coronary blood flow. Physiol. Rev., *44*:1, 1964.
4. Blain, C., Corby, D., Anderson, T., Schulman, I., and Gunnar, R. M.: Bacteremia and shock in man. Clin. Res., *17*:230, 1969.
5. Blumgart, H. L., Schlesinger, M. J., and Davis, D.: Studies in the relationship of the clinical manifestations of angina pectoris, coronary thrombosis and myocardial infarction to pathologic findings. Amer. Heart J., *19*:1, 1940.
6. Cohn, J. N.: Blood pressure measurement in shock. J.A.M.A., *199*:972, 1967.
7. Cohn, J. N., Khatri, I. M., and Hamosh, P.: Diagnostic and therapeutic value of bedside monitoring of left ventricular pressure. Amer. J. Cardiol., *23*:107, 1969.
8. Constantin, L.: Extracardiac factors contributing to hypertension during coronary occlusion. Amer. J. Cardiol., *11*:205, 1963.
9. Edelman, N. H., Hirsch, L., Gorfinkel, J., Gottshalk, A., and Fishman, A. P.: Pulmonary gas exchange in cardiogenic and hypovolemic shock. Clin. Res., *17*:238, 1969.
10. Fluck, D. C., Olsen, E., Pentecost, B. L., Thomas, M., Fillmore, S. J., Shillingford, J. P., and Mounsey, J. P. P.: The natural history and clinical significance of arrhythmias after acute myocardial infarction. Brit. Heart J., *29*:170, 1969.
11. Gilbert, R. P., Aldrich, S. L., and Anderson, L.: Cardiac output in acute myocardial infarction. J. Clin. Invest., *30*:640, 1951.
12. Gilbert, R. P., and Hohf, R.: Hemodynamic basis of norepinephrine shock. Proc. Soc. Exper. Biol. Med., *116*:43, 1964.
13. Gunnar, R. M., Cruz, A., Boswell, J., Co, B. S., Pietras, R. J., and Tobin, J. R., Jr.: Myocardial infarction with shock. Hemodynamic studies and results of therapy. Circulation, *33*:753, 1966.
14. Gunnar, R. M., Loeb, H. S., Pietras, R. J., and Tobin, J. R., Jr.: Hemodynamic measurements in a coronary care unit. Prog. Cardiovasc. Dis., *11*:29, 1968.
15. Gunnar, R. M., Loeb, H. S., Pietras, R. J., and Tobin, J. R., Jr.: Ineffectiveness of isoproterenol in the treatment of shock due to acute myocardial infarction. J.A.M.A., *202*:1124, 1967.
16. Gunnar, R. M., Pietras, R. J., Stavrakos, C., Loeb, H. S., and Tobin, J. R., Jr.: The physiologic basis for treatment of shock associated with myocardial infarction. MED. CLIN. N. AMER., *51*:69, 1967.

17. Guss, J. H.: Chronic shock after myocardial infarction; its mechanism and management. Ann. Int. Med., *58*:333, 1963.
18. Hamilton, W. F., Riley, R. L., Attyah, A. M., Cournand, A., Fowell, D. M., Himmelstein, A., Noble, R. P., Remington, J. W., Richards, D. W., Jr., Wheeler, N. C., and Whitham, A. C.: Fick and the dye injection method of measuring cardiac output in man. Amer. J. Physiol., *153*:309, 1948.
19. Kuhn, L.: The treatment of cardiogenic shock. Amer. Heart J., *74*:578 and 725, 1967.
20. Longsjoen, P. H., Sanchez, S. A., Lynch, D. J., and Inmon, T. W.: The treatment of myocardial infarction with low molecular weight dextran. Amer. Heart J., *76*:28, 1968.
21. Lassers, B. W., Anderton, J. L., George, M., Muir, A. L., and Julian, D. J.: Hemodynamic effects of artificial pacing in complete heart block complicating acute myocardial infarction. Circulation, *38*:308, 1968.
22. Lillehei, R. C., Longerbeam, J. K., Bloch, J. H., and Manax, W. C.: The nature of irreversible shock. Experimental and clinical observations. Ann. Surg., *160*:682, 1964.
23. Lluch, S., Moguilevsky, H. C., Pietra, G., Shaffer, A. B., Hirsch, L. J., and Fishman, A. P.: A reproducible model of cardiogenic shock in the dog. Circulation, *39*:205, 1969.
24. Loeb, H. S., Gunnar, R. M., Pietras, R. J., and Tobin, J. R., Jr.: Relationships between central venous and left ventricular filling pressures prior to and during treatment of shock. Amer. J. Cardiol., *23*:125, 1969.
25. Loeb, H. S., Pietras, R. J., Ninos, N. P., Tobin, J. R., Jr., and Gunnar, R. M.: Hemodynamic responses to chlorpromazine in patients in shock. Arch. Int. Med., *124*:354, 1969.
26. Loeb, H. S., Pietras, R. J., Tobin, J. R., Jr., and Gunnar, R. M.: Hypovolemia in shock due to acute myocardial infarction. Circulation, in press.
27. MacKenzie, G. J., Taylor, S. H., Flenley, D. C., McDonald, A. H., Staunton, H. P., and Donald, K. W.: Circulatory and respiratory studies in myocardial infarction and cardiogenic shock. Lancet, *2*:825, 1964.
28. Montgomery, E. F., Co, B. S., Pietras, R. J., Gunnar, R. M., and Tobin, J. R., Jr.: Immediate and delayed hemodynamic changes resulting from the restoration of atrial systole by electroversion. Circulation, *34*, Suppl. III, 1966, p. 172.
29. Morse, B. W., Danzig, R., and Swan, H. J. C.: Effect of isoproterenol in shock associated with acute myocardial infarction. Circulation, *36*, Suppl. 2, 1967, p. 192.
30. Mosher, P., Ross, J., McFate, R. A., and Shaw, R. F.: Control of coronary blood flow by an autoregulatory mechanism. Circ. Res., *14*:250, 1964.
31. Oriol, A., and McGregor, M.: Indicator dilution methods in estimation of cardiac output in clinical shock. Amer. J. Cardiol., *20*:826, 1967.
32. Rapaport, E., *In* Proceedings of Symposium on Research on Acute Myocardial Infarction, American Heart Association, Phoenix, Arizona, March 26, 1969.
33. Shapiro, M., Fillmore, S., and Killip, T.: Hypoxemia and functional severity in acute myocardial infarction. Circulation, *38*, Suppl. 6, 1968, p. 178.
34. Smith, H. J., Oriol, A., March, J., and McGregor, M.: Hemodynamic studies in cardiogenic shock: Treatment with isoproterenol and metaraminol. Circulation, *35*:1084, 1967.
35. Smith, J., Pennincky, J. J., Kampschulte, S., and Safar, P.: Need for oxygen enrichment in myocardial infarction, shock and following cardiac arrest. Acta. Anesth. Scandinav., Suppl. *29*, 1968, p. 127.
36. Talley, R. C., Goldberg, L. I., Johnson, C. E., and McNay, J. L.: A hemodynamic comparison of dopamine and isoproterenol in patients in shock. Circulation, *39*:361, 1969.
37. Weil, M. H., Shubin, H., and Rosoff, L.: Fluid repletion in circulatory shock: Central venous pressure and other practical guides. J.A.M.A., *192*:668, 1965.
38. Wiggers, C. J.: The functional consequences of coronary occlusion. Ann. Int. Med., *23*: 158, 1945.
39. Wiggers, C. J., Wegria, R., and Pinera, B.: The effects of myocardial ischemia on the fibrillation threshold. The mechanism of spontaneous fibrillation following coronary occlusion. Amer. J. Physiol., *131*:309, 1940.

840 S. Wood Street
Chicago, Illinois 60612

Electrical Slowing of the Heart and Post-Extrasystolic Potentiation

LEON RESNEKOV, M.D.(Cape Town), M.R.C.P.*

Electrical stimulation of the heart has been attempted ever since Aldini[1] recommended galvanic current for treating syncope. Gould[13] successfully resuscitated a baby using a needle electrode directly into the myocardium; Zoll[31] was able to stimulate the heart through the chest wall, and ultimately Hopps[21] demonstrated the feasibility of pacing the atrium using a transvenous electrode. While electrical methods of increasing the heart rate were perfected in this way, there was still no satisfactory method for slowing an inappropriately rapid heart rate until in 1963 Lopez, Edelist, and Katz[25] were performing an experiment on a dog using a particularly long pulse for electrical stimulation. To their surprise and delight, this long pulse succeeded in almost halving the heart rate (Fig. 1). They were able to show that this effect was related to the heart being stimulated at both the "make" and "break" of the wide pulse wave, and that a similar effect could be produced by using two short stimuli separated by the duration of the original long pulse wave. When two such stimuli are delivered to the ventricle, the technique is called *paired pulse* stimulation, and earlier reports[6, 4, 8, 27] confirmed its use in slowing the heart.

The essential feature of the technique is that the first pulse overdrives the heart and causes a mechanical contraction; the second is timed to occur in the relative refractory period, causes electrical depolarization unaccompanied by mechanical contraction, prolongs the refractory period, and reduces the ventricular rate. The interval between the two pulses varies according to the heart rate, but in humans is usually in the order of 300 msec.

In contrast, when *coupled pacing*[10] is employed, a single electrical stimulus is applied at the appropriate interval following an R-wave of a

*Associate Professor of Medicine (Cardiology), Department of Medicine, University of Chicago, Chicago, Illinois

Supported in part by grants from the National Heart Institute, National Institutes of Health (He-05673-05); Myocardial Infarction Research Unit (contract PH 4368133); and the Chicago and Illinois Heart Associations.

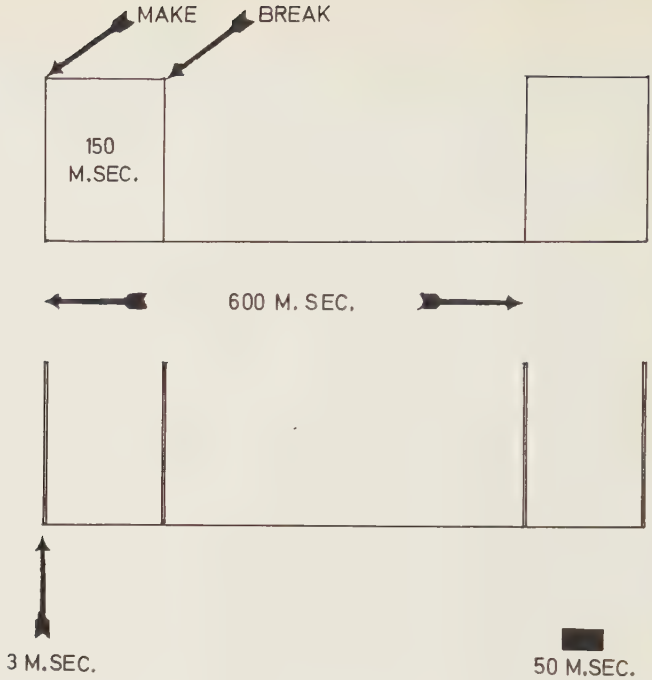


Figure 1. Original experiment performed by Lopez, Edelist, and Katz,[25] in which a stimulating pulse of 150 msec. duration was used (upper part). The heart rate was slowed due to stimulation by both the "make" and "break" (see text). A similar effect was produced by stimulating the heart using two 3 msec. pulses separated by the duration of the long pulse (lower part).

spontaneous ventricular contraction. Once more, prolongation of the refractory period of the ventricle is achieved; the next atrial impulse arrives during the refractory phase and does not result in ventricular contraction. An immediate halving of the heart rate is therefore achieved. The difference between paired and coupled pacing is that with coupled pacing, the heart rate is not controlled by the stimulator, the normal atrioventricular relationships are preserved, and the benefit of atrial systole is maintained. The experimental and clinical applications of this technique have been well documented.[4, 5, 10]

In the presence of normal atrioventricular conduction, paired stimulation can be applied to the atrium. The second of each pair of atrial depolarizations fails to pass through the atrioventricular node, since it is still refractory from the first. Slowing of the ventricular rate is thereby achieved, and the atrial contribution to ventricular filling is still maintained.

This communication is concerned primarily with paired stimulation of the ventricle, during which the first stimulus drives the ventricle at a fixed rate to cause an effective mechanical systole; the second, timed to occur in the relative refractory period, results in an electrical premature beat without any corresponding mechanical contraction; the refractory

period is prolonged and the heart rate slowed. A premature electrical depolarization precedes each ventricular contraction, and therefore every mechanical beat produced by the technique shows the phenomenon of post-extrasystolic potentiation.[18]

METHOD

A stimulator was designed and built and used to investigate the circulatory effects and clinical indications of paired stimulation of the ventricle (Fig. 2). Using this apparatus, stimuli of 1 msec. duration can be delivered singly or in pairs at rates between 30 and 250 per minute and at up to 10 V. Each stimulus can be varied independently, and the delay between the stimuli can be adjusted between 35 and 350 msec. Experimental studies were performed in dogs under general anesthesia with sodium pentobarbital during maintained respiration with room air delivered through a cuffed endotracheal tube from a positive pressure pump. A left thoracotomy was used and the heart was stimulated directly from the atrium or from the ventricle.

Left ventricular pressure was measured through a short length of Teflon tubing introduced into the apex of that chamber. The electrocardiogram was recorded from a right atrial lead; total blood flow immediately beyond the origin of the coronary arteries was measured using an electromagnetic flowmeter; systemic arterial pressure was measured from the ascending aorta; and the first derivative of the ventricular pressure pulse was derived. All these parameters were recorded simultaneously on a multichannel recording system. The clinical application of the technique was studied in patients seriously ill as a result of resistant rhythm disturbances and also in patients with intractable heart failure

Figure 2. External pacemaker for single or paired stimulation.

to determine whether electroaugmentation[28] would occur and result in improvement of cardiovascular function.

RESULTS

Animal Experiments

HEART RATE. *Sinus rhythm.* Slowing the heart below its spontaneous sinus rate was achieved on all occasions, the maximal reduction being 45 per cent of the control value (Fig. 3). Sinus tachycardia was easily slowed with two impulses separated by a delay of 150 to 250 msec. It was often necessary to readjust this delay, nor did it remain constant over prolonged periods of time; a hump appearing on the downslope of the left ventricular pressure curve indicated that it was too long, allowing the ventricle to contract partially to the second stimulus; as the interval was shortened, this hump disappeared. When the delay was shortened even further, however, the second stimulus of the pair moved into the absolute refractory period so that the effect of paired stimulation was lost and a sinus beat interposed (Fig. 4).

Supraventricular tachycardia. An artificially created supraventricular tachycardia was produced by driving the atrium with a second pacemaker at rates of up to 250 beats per minute. This tachycardia was consistently slowed by paired stimulation of the ventricle, when rates of 216 beats per minute could be easily reduced to 120 beats per minute (Fig. 5).

Ventricular tachycardia. In a similar way ventricular tachycardia was produced by artificially stimulating the ventricle directly using an intramural electrode and a second pacemaker (Fig. 6). A profound fall in aortic flow and blood pressure usually resulted. Paired stimulation of the ventricle was always effective in reducing the ventricular contraction rate, during which time both the pressure and the flow were increased. The ventricular ectopic focus producing the tachycardia could not be depolarized, since the second driving pacemaker continued to function, and interference cycles between the pacemaker and the paired stimulator were therefore recorded. Breakthrough appeared when impulses from the driving pacemaker produced short runs of ventricular tachycardia associated with a fall in pressure and flow (Fig. 7). Nevertheless, on balance there was considerable over-all improvement and the average ventricular rate was slowed. Even under the worst possible conditions, therefore, when the ectopic focus could not be depolarized, the effects of the ventricular tachycardia were still diminished by paired stimulation.

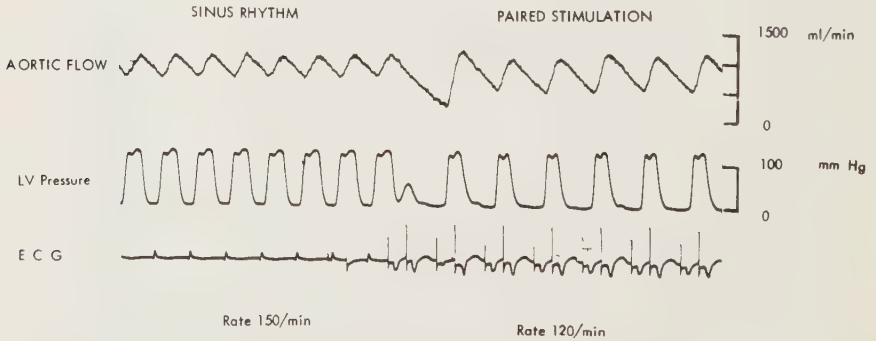


Figure 3. Sinus rhythm slowed from 150 to 120 beats per minute by paired stimulation of the ventricle. Aortic flow, LV pressure, and the electrocardiogram recorded from above downwards. Paired stimulation causes two electrical depolarizations but only one mechanical beat for each pair. Note slurring of the descending limb of the ventricular pressure pulse during paired stimulation.

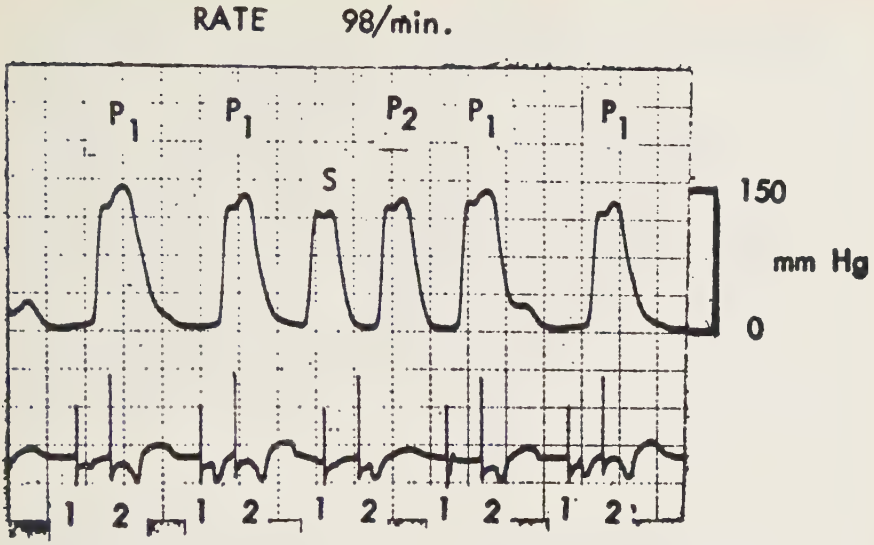


Figure 4. Incorrect adjustment of the delay between first and second stimuli. Paired stimulation has resulted in an effective ventricular rate of 98/min. (above). The driving stimulus of the ECG (below) is labeled 1 and the paired stimulus 2. The first two ventricular beats are caused by the first stimulus of each pair and are labeled P_1. The delay of the paired stimulus of the second beat is short, the stimulus occurs in the absolute refractory period and a sinus beat (5) has interposed. The beat following (P_2) has resulted from the second (2) of the next pair of stimuli, for the ventricle was refractory to the first (1), but thereafter satisfactory fusion is shown.

ELECTROAUGMENTATION. Changes in the first derivative of the left ventricular pressure pulse (dP/dt) are often used an an index of myocardial potentiation, but their dependence on heart rate must be allowed for.[27] It is important, therefore, that electroaugmentation during paired stimulation be demonstrated at a rate identical to the control sequence. When this is done, a significant increase

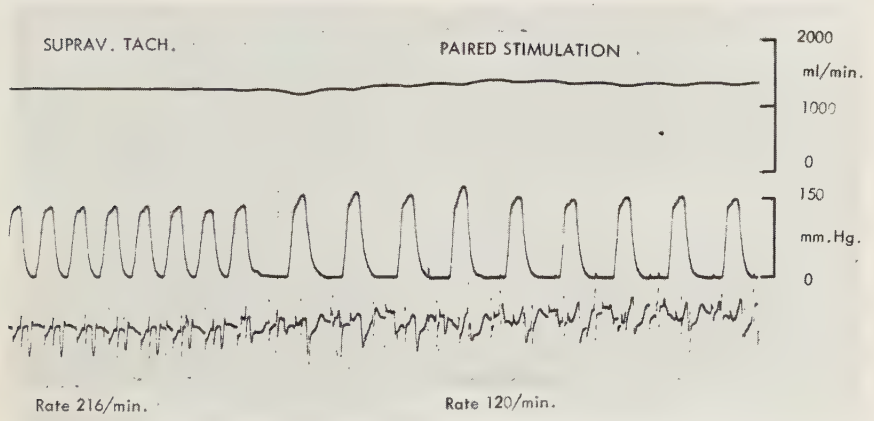


Figure 5. Slowing of a supraventricular tachycardia (pacemaker-induced). Mean flow in the ascending aorta (above) recorded with LV pressure (middle) and RA electrogram (below). Paired stimulation of the ventricle has slowed the rate from 216 to 120 beats/min. and is associated with an increase in pressure and a slight increase in flow.

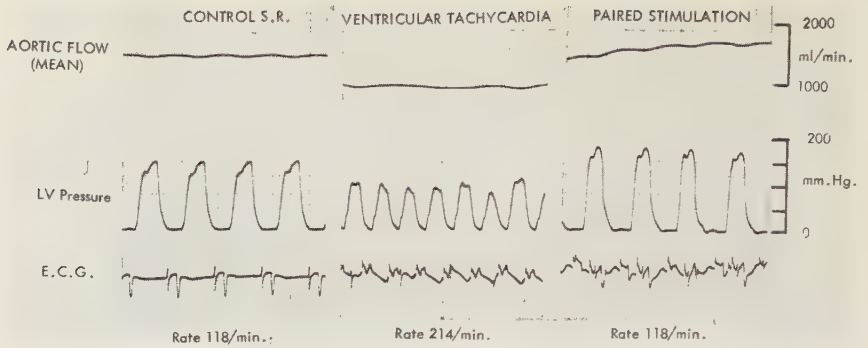


Figure 6. Paired stimulation of pacemaker-induced ventricular tachycardia. The layout is the same as in Figure 5. Ventricular tachycardia at 214 beats/min. is associated with a fall in pressure and flow, both of which are corrected by slowing the rate to 118/min. by paired stimulation.

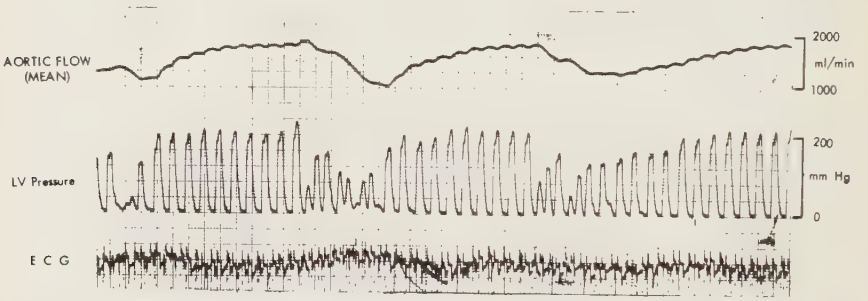


Figure 7. Paired stimulation of the pacemaker-induced ventricular tachycardia. Interference cycles between the driving pacemaker and the paired stimulator are associated with short runs of ventricular tachycardia and a reduction of LV pressure and aortic flow.

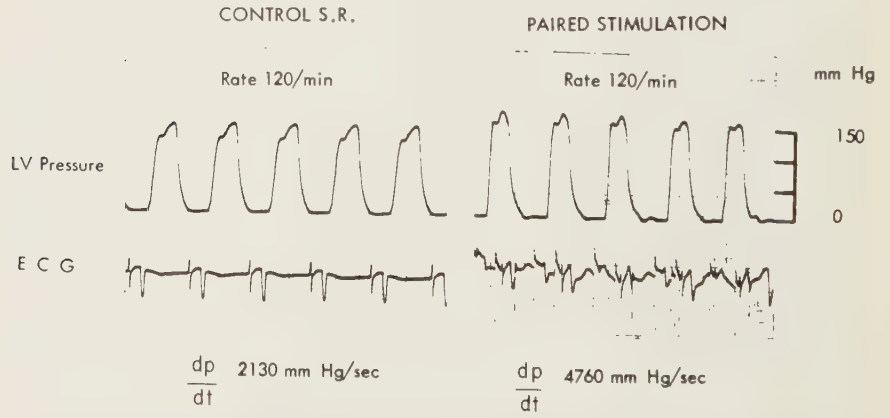


Figure 8. Electroaugmentation produced by paired stimulation of the ventricle. During paired stimulation at the same rate as the control (120/min.) LV pressure has increased by 16 per cent and dP/dt by 125 per cent.

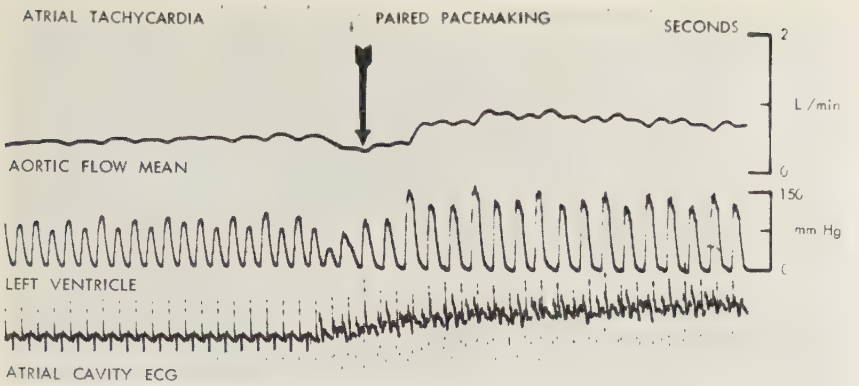


Figure 9. Positive inotropic action of paired stimulation. A paced atrial tachycardia at 190 beats/min. is slowed to 140/min. during paired stimulation of the ventricle. An increase in mean aortic flow and LV systolic pressure are recorded and a fall in LV end-diastolic pressure.

in both left ventricular pressure and in dP/dt is demonstrated at equivalent ventricular rates (Fig. 8), despite the absence of atrial contribution to ventricular filling during paired stimulation, since the stimulating electrodes are attached to the ventricle. When cardiac output was reduced during the control period, a sustained increase in flow was always recorded during paired stimulation (Fig. 9), but no increase in flow could be demonstrated unless this reduction during the control period was significant. In addition, a further positive inotropic effect was shown by a decrease in the left ventricular end-diastolic pressure when that pressure was raised.

Clinical Use

Five illustrative examples are given of the use of paired ventricular stimulation clinical practice, both for the control of intractable rhythm disturbances and in an attempt to improve resistant cardiac failure:

CASE 1. A man, 23 years old, known to have the Wolff-Parkinson-White syndrome, developed atrial tachycardia at a ventricular rate of 180 per minute, which was completely resistant to drug therapy and electroversion. Severe heart failure occurred, extreme peripheral vasoconstriction was present, and he became semiconscious due to low cardiac output. Paired stimulation of the ventricle was attempted following positioning of an electrode catheter in the apex of the right ventricle. The heart responded to both stimuli and ventricular fibrillation occurred immediately thereafter. Fortunately, this was easily converted to sinus rhythm by a single unsynchronized direct current shock.

CASE 2. A woman 58 years old was admitted with ventricular tachycardia and a heart rate of 170 per minute in association with recent myocardial infarction. The dysrhythmia was resistant to treatment, and a decision was made to attempt control by means of paired stimulation of the right ventricle. It was impossible to find a satisfactory delay between the two stimuli, and each attempt at paired stimulation provoked runs of ventricular extrasystoles despite the fact that the stimulus strength was reduced to just above threshold levels. Satisfactory fusion could not be produced and the attempt was therefore abandoned.

CASE 3. A man of 50 years developed resistant supraventricular tachycardia alternating with atrial fibrillation in the early postoperative period following aortic and mitral valve replacement. The dysrhythmias were completely resistant to drug therapy and electrolyte correction. During the tachycardias, the pulmonary arterial saturation was 34 per cent and the arteriovenous oxygen difference 83

ml. per liter. Paired stimulation of the right ventricle was undertaken and easily maintained for 48 hours. The heart rate was successfully reduced from 220 to 140 per minute. During this time the pulmonary arterial saturation rose to 56 per cent and the arteriovenous oxygen difference fell to 56 ml. per liter, indicating an increase in cardiac output of some 50 per cent over the control value.

CASE 4. A man of 64 years, known to have had atrial fibrillation for some time prior to his present admission with myocardial infarction, developed severe cardiac failure associated with rupture of the ventricular septum. Heart failure was resistant to treatment, and in addition was associated with severe and progressive renal dysfunction. Paired stimulation of the ventricle was decided on to obtain the benefit of electroaugmentation and improve his circulatory state before surgical repair of the ventricular septum. Cardiac catheterization confirmed a left-to-right shunt at ventricular level and demonstrated a pulmonary:systemic flow ratio of 2:1. In addition, moderate pulmonary hypertension secondary to the effect of the shunt and a greatly raised end-diastolic pressure in the left ventricle were recorded. Severe tricuspid regurgitation was also present. During paired stimulation of the ventricle, however, there was no significant fall in the end-diastolic pressure of the left ventricle, nor were any beneficial changes recorded in the aortic, pulmonary arterial, or right atrial pressures. Electroaugmentation did not occur, and the degree of left-to-right shunt remained unchanged. Although some improvement in cardiac output and stroke volume was noted, this change was entirely rate-dependent, and there was no evidence that a positive inotropic effect was obtained. The technique was maintained for 24 hours without objective clinical or hemodynamic evidence of benefit. Thereafter the patient was referred for surgery, which was successfully performed.

CASE 5. A woman of 48 years was admitted in severe chronic left and right ventricular failure secondary to mitral stenosis and regurgitation, aortic stenosis and regurgitation, restrictive pulmonary hypertension, tricuspid regurgitation, and atrial fibrillation. The control ventricular rate during atrial fibrillation was 80 per minute, and paired stimulation was attempted to improve myocardial function prior to mitral and aortic valve replacement. Paired stimulation at a similar ventricular rate produced a marked deterioration. The cardiac output fell by 20 per cent, the gradient across the mitral valve rose by some 60 per cent, and a profound fall in systemic arterial pressure was recorded. Electroaugmentation did not occur. The hemodynamic state was considerably worsened, and the attempt at paired stimulation had to be abandoned.

DISCUSSION

During paired stimulation of the ventricle,[25] the first of the pair of electrical stimuli drives the ventricle at a fixed rate, and the second, which is timed to occur in the relative refractory period, causes electrical depolarization unaccompanied by mechanical contraction. In essence, there are two electrical events for each mechanical contraction, the refractory period is prolonged, and the heart rate is slowed. The sinus node may be suppressed during the technique or sinoatrial dissociation may occur, but it is more likely that the atrioventricular node is kept unresponsive by concealed retrograde conduction, although retrograde stimulation of the atria may also be caused.

If premature electrical depolarization precedes each mechanical contraction, the phenomenon of post-extrasystolic potentiation will occur.[18] This effect is an inherent property of mammalian cardiac muscle, and although in the intact heart there is a longer time for filling of the ventricle during the compensatory pause which follows an extra-

systolic beat, this is not the only reason for it to occur; potentiation follows an interpolated extrasystole in a similar fashion,[12] and in addition, isolated strips of atrial and ventricular muscle mounted in such a way that only isometric contraction occurs also show the effect. The earlier the post-extrasystolic beat, the greater is the amount of potentiation caused.[30]

Augmentation will still occur when the extrasystole is so early in the relative refractory period that no mechanical response is possible.[23] The effect is greatest in a ventricle which contracts against a heavy resistive load or a failed ventricle.[7, 19] An increase in the maximal force of contraction during systole to 64 per cent of the control value at any ventricular volume can be shown, and there is an increase in the rate of development of tension, an increase in the rate of relaxation, and an increase in the velocity of myocardial fiber shortening.[11] When myocardial failure is present, a sustained positive inotropic action can be seen by a reduction in the ventricular end-diastolic pressure when that pressure is raised, and increase in the cardiac output (when that is reduced), and a decrease in the size of the heart (when it is enlarged).

Most investigators feel that the reduction in the ventricular end-diastolic pressure produced by paired stimulation is due to an associated increase in diastolic compliance of the heart.[2] If this is so, it follows that the end-diastolic pressure can no longer be accepted as a guide to end-diastolic volume.[7] Although the effects of post-extrasystolic potentiation have been clearly demonstrated, the fundamental mechanism underlying the phenomenon is not known. Its effect can always be demonstrated unless levels of calcium are very high,[26] nor does the local release of norepinephrine play a significant role, for potentiation occurs even when hearts have been treated with reserpine.[28] It is likely that the loss of potassium ions may be important[15] and that an influx of calcium ions related to the degree of depolarization might be a cause of the positive inotropic effects produced.[14]

The effect of the technique in reversing acute heart failure in experimental animals is dramatic and gratifying. Slowing of the heart results in a longer diastolic filling time and a bigger stroke volume, but the cardiac output is not increased unless myocardial function is depressed during the control period. This positive inotropic action is most marked when myocardial function is more severely affected, possibly due to the differing contractility response of the failing and nonfailing heart.[11] Paired stimulation is similar in its effects to digitalis, although acetyl strophanthidine tends to elevate arterial pressure more than does paired stimulation.[11] The increase in ventricular contractility produced by paired stimulation in the intact heart may still be augmented by an infusion of norepinephrine,[9] unlike the effect of paired stimulation on isolated heart muscle preparations.

Despite all the articles written on the experimental and basic aspects of the technique, indications for its clinical use are not yet established. This is regrettable, for its action can be most useful, particularly in controlling resistant dysrhythmias, and it should be available in every intensive care unit. It is as well to remember, however, that it is a

demanding technique and should be employed only by those trained in its application, for patients treated by paired stimulation require constant attention and frequent adjustment of the delay between the two stimuli. The five clinical examples presented in this paper were chosen to demonstrate not only the beneficial effects of the technique, but also what problems and difficulties might be encountered when it is used. It is most important to begin stimulation in such a way that the heart will not respond to both electrical pulses by determining first the appropriate driving rate of the first stimulus and then the correct delay between the two; small adjustments may then be made safely to obtain the maximal potentiating effect if desired.

The following method for induction is recommended. The stimulator rate is set to just above half the spontaneous rate of the heart, the interval between the two pulses is set deliberately too long (350 msec.), and the stimulus strength is increased slowly until the heart responds to every stimulus. The interval between the two stimuli is then progressively reduced until fusion occurs, when the heart rate will suddenly halve and electroaugmentation will be produced. Although not essential, a simultaneous record of a ventricular pressure pulse is desirable and helpful in adjusting the delay; the stimuli should be at the lowest effective electrical energy by using an impulse of not more than 1 msec. width at just above threshold levels to lessen the risk of ventricular fibrillation. The technique for discontinuing paired stimulation is to allow the spontaneous pacemaker to dominate rapidly over several beats by stopping the second impulse and reducing the rate and amplitude of the first.

Even when all precautions are observed, certain groups of patients may be unsuitable for paired stimulation. Our experience with the technique in patients with known coronary arterial disease has been unhappy. The second impulse has to be placed close to the vulnerable phase of the ventricle for fibrillation (Fig. 10), but this by itself is not likely to

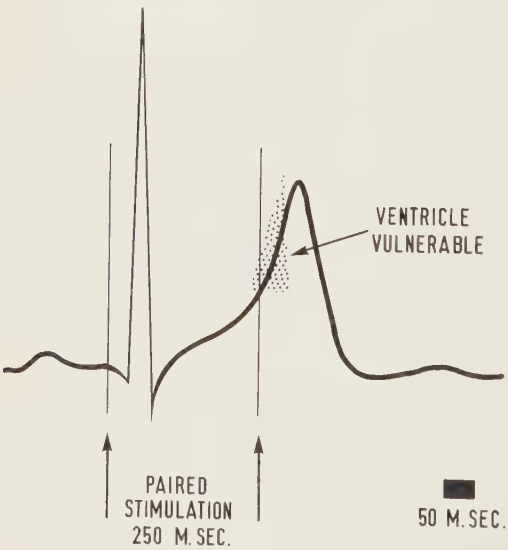


Figure 10. Vulnerable phase of the ventricle for fibrillation, showing the relating positions of the driving and pairing stimuli.

be the cause. Paired stimulation following acute experimental coronary arterial occlusion results more frequently in ventricular fibrillation than does single pacemaking,[17] and a slower heart rate is implicated.[16] The slower rhythm results in asynchrony of recovery of excitability of the ventricles. Similarly, irregular perfusion of the ventricles will also predispose to ectopic rhythms.[16]

There is a risk of the induced post-extrasystolic beat causing repetitive activation when the technique is used in patients with coronary arterial disease or atrioventricular dissociation.[7] Under these circumstances, a relatively high basic heart rate should always be chosen, and the threshold for stimulation kept at no more than twice the diastolic level. Many investigators now feel that coronary arterial disease and conduction disturbances are both relative contraindications to using paired stimulation of the ventricle.

The technique has had its most useful effect in the early postoperative phase following cardiopulmonary bypass, particularly in the control of resistant rapid heart rates, and also for its positive inotropic effect. It should be considered in any patient in whom electroversion and drug therapy fail to control the heart rate or rhythm and should be used before serious circulatory effects of the tachycardia are allowed to develop. Its use in acute or chronic heart failure to improve myocardial function where rate alone is not the problem is much more debatable. Our experience under these circumstances has been far from dramatic, and in the two examples presented in this paper (Cases 4 and 5), the loss of atrial systole accompanying the use of paired stimulation could not be blamed for the lack of a positive inotropic effect; in both instances atrial fibrillation was present during the control state.

Despite experimental evidence to the contrary, clinical experience indicates that the increased myocardial contractility which follows paired stimulation is not often accompanied by an increase in cardiac output when paired stimulation is used in patients with reduced myocardial function. Fusion is never quite complete; slurring of the descending limb of the ventricular pressure pulse is almost inevitably seen and may impede ventricular filling. Furthermore, the pathway of ventricular excitation is quite abnormal, and closure of the atrioventricular valves may be interfered with.[29] It may well be that where it is important to preserve atrial systole, coupled rather than paired stimulation (in the absence of atrial fibrillation) should be used, or paired stimulation of the atrium (in the absence of atrial fibrillation or ventricular tachycardia) should be attempted. It should be appreciated, however, that when paired stimulation of the atria is employed, rate control may be achieved but electroaugmentation is not produced.

Severe hemodynamic deterioration followed the use of paired stimulation in patient 5, who had severe mitral and tricuspid regurgitation, and severe atrioventricular valve regurgitation may also be a contraindication. Myocardial oxygen consumption is increased during paired stimulation of the ventricle,[4, 5, 8] even when the cardiac output remains unchanged. Increases in the order of 35 per cent are produced, resulting particularly from the increased speed of contraction[28] and not from the

second electrical depolarization, which could only be responsible for a 1 per cent increase.[22] Paired stimulation of the atria, on the other hand, does not demand a concomitant increase in myocardial oxygen consumption,[24] presumably because of the absence of any associated electroaugmentation.

The inability to provide the additional myocardial oxygen requirement because of coronary arterial disease or severe ventricular hypertrophy might well be responsible for the disappointing results observed when paired stimulation of the ventricle is used to improve myocardial function. Triple stimulation of the ventricle, although occasionally useful in controlling rhythm disturbances not responding to paired stimulation,[3] is as limited as paired stimulation in this respect; additional electroaugmentation is not obtained.[10] The hope that during paired stimulation an associated increase in coronary arterial flow[20] would compensate for any increased myocardial oxygen consumption is unfortunately not often realized in its clinical use.

SUMMARY

The basic principles of paired stimulation of the ventricle are described and the techniques of coupled ventricular stimulation and paired stimulation of the atrium detailed. Circulatory effects of paired stimulation in experimental animals are reported and its effect clinically described. Although demanding as a technique, paired stimulation of the ventricle should be available in any unit undertaking intensive care, particularly for its action in controlling ventricular rate; dysrhythmias in the postoperative period following open heart surgery respond particularly well. Undue slowing of the heart is undesirable in patients with coronary arterial disease, in whom the technique has to be used with extreme caution, and patients with conduction defects may also be unsuitable for paired stimulation. Atrioventricular regurgitation is increased during paired stimulation of the ventricle, and the technique is used with caution, if at all, in such patients. Its clinical use to improve myocardial function when rate disturbance is not present is often disappointing, and the reasons for this are discussed with reference to underlying mechanisms and myocardial oxygen requirements. Coupled rather than paired stimulation of the ventricle or paired stimulation of the atrium in the absence of conduction defects, atrial fibrillation, or ventricular tachycardia might be a more appropriate technique in the presence of myocardial dysfunction.

ACKNOWLEDGMENT

Figures 3 to 8 are reproduced with the permission of the Editor, British Medical Journal.

REFERENCES

1. Aldini, G.: General views on the application of galvanism to medical purposes principally in cases of suspended animation. London, Callow, 1819.
2. Bartelstone, H. J. B., Scherlag, B. J., Hoffman, B. F., and Cranefield, P. F.: Demonstration of variable diastolic compliance associated with paired stimulation of the dog heart. Bull. N.Y. Acad. Med., *41*:616, 1965.

3. Bayley, T. J., and Lightwood, R.: Double and triple pulse pacemaking in treatment of ventricular tachycardia. Lancet, *1*:235, 1966.
4. Braunwald, E., Ross, J., Jr., Frommer, P. L., Williams, J. F., Jr., Sonnenblick, E. H., and Gault, J. H.: Clinical observations on paired electrical stimulation of the heart. Amer. J. Med., *37*:700, 1964.
5. Braunwald, E., Ross, J., Jr., Sonnenblick, E. H., Frommer, P. L., Braunwald, N. S., and Morrow, A. G.: Slowing of heart rate; electroaugmentation of ventricular performance and increase of myocardial oxygen consumption produced by paired electrical stimulation. Bull. N.Y. Acad. Med., *41*:481, 1965.
6. Chardack, W. M., Gage, A. A., and Dean, D. C.: Slowing of the heart by paired pulse pacemaking. Amer. J. Cardiol., *14*:374, 1964.
7. Cranefield, P. F.: Paired pulse stimulation and post-extrasystolic potentiation in the heart. Progr. Cardiovasc. Dis., *8*:446, 1966.
8. Dean, D., Chardack, W., and Gage, A.: Slowing of the heart by paired electrical stimuli. Clin. Res., *12*:180, 1964 (abstract).
9. Fisher, V. J., Lee, R. J., Marlon, A. M., and Kavaler, F.: Paired electrical stimulation and the maximal contractile response of the ventricle. Circ. Res., *20*:533, 1967.
10. Frommer, P. L.: Studies on coupled pacing technique and some comments on paired electrical stimulation. Bull. N.Y. Acad. Med., *41*:670, 1965.
11. Frommer, P. L., Robinson, B. F., and Braunwald, E.: Paired electrical stimulation. A comparison of the effects on performance of the failing and non-failing heart. Amer. J. Cardiol., *18*:738, 1966.
12. Garb, S., and Penna, M.: Some quantitative aspects of the relation of rhythm to the contractile force of mammalian ventricular muscle. Amer. J. Physiol., *182*:601, 1965.
13. Gould, A. S., 1929, as quoted by Hyman, A. S.: Resuscitation of the stopped heart by intracardial therapy. Arch. Int. Med., *50*:283, 1932.
14. Grossman, A., and Furchgott, R. F.: The effects of frequency of stimulation and calcium concentration on Ca^{45} exchange and contractility on the isolated guinea pig auricle. J. Pharmacol. Exper. Therap., *143*:120, 1964.
15. Hajdu, S., and Leonard, E.: The cellular basis of cardiac glycoside action. Pharmacol. Rev., *11*:173, 1959.
16. Han, J., Millet, D., Chizzonitti, B., and Moe, G. K.: Temporal dispersion of recovery of excitability in atrium and ventricle as a function of heart rate. Amer. Heart J., *71*:481, 1966.
17. Han, J., De Traglia, J., Millet, D., and Moe, G. K.: Incidence of ectopic beats as a function of basic rate in the ventricle. Amer. Heart J., *72*:632, 1966.
18. Hoffman, B. F., Bindler, G., and Suckling, E. E.: Post-extrasystolic potentiation of contraction in cardiac muscle. Amer. J. Physiol., *185*:95, 1956.
19. Hoffman, B. F., Bartelstone, H. J., Scherlag, B. J., and Cranefield, P. F.: Effects of post-extrasystolic potentiation on normal and failing hearts. Bull. N.Y. Acad. Med., *41*:498, 1965.
20. Hoffman, B. F., and Cranefield, P. F.: Paired-pulse stimulation of the heart. Amer. Heart J., *71*:143, 1966.
21. Hopps, J. A., and Bigelow, W. G.: Electrical treatment of cardiac arrest: a cardiac stimulator-defibrillator. Surgery, *36*:833, 1954.
22. Klocke, F. J., Braunwald, E., and Ross, J., Jr.: The oxygen cost of electrical activation of the heart. Clin. Res., *13*:526, 1965 (abstract).
23. Koch-Weser, J., and Blinks, J. R.: The influence on the interval between beats on myocardial contractility. Pharmacol. Rev., *15*:601, 1963.
24. Lister, J. W., Damato, A. N., Kosowsky, B. D., Lau, S. H., and Stein, E.: The hemodynamic effect of slowing the heart rate by paired or coupled stimulation of the atria. Amer. Heart J., *73*:362, 1967.
25. Lopez, J. F., Edelist, A., and Katz, L. N.: Slowing of the heart rate by artificial electrical stimulation with pulses of long duration in the dog. Circulation, *28*:759, 1963 (abstract).
26. Meijler, F. L., and Durrer, D.: Physiological and clinical aspects of paired stimulation. Bull. N.Y. Acad. Med., *41*:575, 1965.
27. Resnekov, L., Sowton, E., Lord, P., and Norman, J.: Haemodynamic and clinical effects of paired stimulation of the heart. Brit. Heart J., *28*:622, 1966.
28. Ross, J., Jr., Sonnenblick, E. H., Kaiser, G. A., Frommer, P. L., and Braunwald, E.: Electroaugmentation of ventricular performance and oxygen consumption by repetitive application of paired electrical stimuli. Circ. Res., *16*:325, 1965.
29. Sarnoff, S. J., Gilmore, J. P., and Mitchell, J. H.: Influence of atrial contraction and relaxation on closure of the mitral valve. Circ. Res., *11*:26, 1962.
30. Siebens, A. A., Hoffman, B. F., Cranefield, P. F., and Brooks, McC. C.: Regulation of contractile force during ventricular arrhythmias. Amer. J. Physiol., *197*:968, 1959.
31. Zoll, P. M.: Resuscitation of the heart in ventricular standstill by external electrical stimulation. New Eng. J. Med., *247*:768, 1952.

950 East 59th Street
Chicago, Illinois 60637

The Many Forms of Mitral Incompetence

GUSTAVO BERMUDEZ, M.D.*

K. G. NAIR, M.D., Ph.D.**

HANS H. HECHT, M.D.***

Symptomatic manifestations of mitral valve disease were recognized more than 250 years ago by Vieussens.[34] Corvisart, Napoleon's physician, gave the first description of the abnormalities of mitral valve leaflets in his "Essai sur les maladies du Coeur."[5] Hope later distinguished stenosis from incompetence of the mitral valve and pointed to the systolic timing of the mitral regurgitant murmur.[13] Much later, in 1892, Griffith suggested that a *late* apical systolic murmur might also be due to mitral regurgitation.[11] However, this view did not gain ground for nearly three decades because of the teachings of Lewis,[18] Mackenzie,[19] and Cabot,[3] who maintained that systolic murmurs without any other evidence of heart disease were "innocent" in nature. In fact, murmurs confined to *late* systole were considered "innocent" or "functional" until a few years ago.

The pathologic significance of many otherwise often asymptomatic systolic apical murmurs was rediscovered by Sprague and White[32] and by Freeman and Levine,[9] while a correlation of hemodynamics and clinical findings of mitral valve disease was provided by Paul Wood in his classical essay on mitral stenosis.[35] However, a clear recognition of the different types of mitral insufficiency has occurred only recently in the wake of the extensive experience during open heart surgery. Condi-

From the Department of Medicine (Cardiology), University of Chicago, Chicago, Illinois

*Research Associate (Instructor)

**Formerly Assistant Professor of Medicine and Physiology; presently Associate Professor of Medicine, University of Pennsylvania School of Medicine, Philadelphia, Pennsylvania

***Professor of Medicine and Physiology

Supported in part by grants from the United States Public Health Service (HE-05793-03 and PH-43-68-1334) and from the Chicago and Illinois Heart Associations (N 68-70 and C 68-24).

tions that give rise to the late systolic murmur have been delineated with greater accuracy than in the past and the "wastepaper basket" diagnosis of "functional murmur" need no longer be applied to nonrheumatic mitral regurgitation.

In our classification of the different forms of mitral insufficiency (Table 1), we have attempted to define the conditions according to the anatomic basis of valvar dysfunction. Some of these forms of mitral insufficiency, particularly the nonrheumatic variety, are illustrated with case reports.

CASE 1. Papillary Muscle Dysfunction (coronary artery disease) (Table 1, IIIa1).

J.P., a 65-year-old man, was well until November 1967, when he developed progressive breathlessness on effort and fatigability. In January 1968 an apical systolic murmur was heard for the first time. There was no previous history of myocardial infarction, rheumatic fever, streptococcal infection, trauma, or subacute bacterial endocarditis. On examination the jugular venous pressure was moderately elevated. The radial pulse was regular, 60 per minute, and of normal quality. The blood pressure was 150 mm. Hg systolic, 90 diastolic. There was a

Table 1. *Causes of Mitral Incompetence*

I. *Valvar Lesions*
   a. Rheumatic heart disease (RHD) (Case 4, Fig. 4)
   b. Bacterial endocarditis
   c. Congenital lesions: cleft valve; maldeveloped cusp; "Ebstein's malformation" in corrected transposition of the great vessels; double mitral orifice; anomalous mitral arcade; fused fixed valve; absence of commissural leaflet tissue (Case 7, Fig. 7)
   d. Myxomatous transformation of the valve (floppy valve syndrome) (Case 6, Fig. 6)
   e. Systemic lupus erythematosus (Libman-Sacks syndrome)
   f. Trauma: surgical or nonsurgical
   g. Tumors interfering with valvar closure
   h. Failure of prothesis or homograft

II. *Annular Lesions*
   a. Annular dilatation in chronic congestive failure
   b. Mitral annular calcification (Rytand) (Case 5, Fig. 5)
   c. Annular calcification secondary to subvalvar aneurysm (in Bantus)
   d. Annular dilatation due to left atrial myxoma

III. *Subvalvar Lesions* (subvalvar mitral incompetence; SMI)
   a. *Papillary muscle dysfunction*
      1. Infarction, with or without rupture (Case 1, Fig. 1)
      2. Anomalous left coronary artery with fibrosis of papillary muscle (Case 8)
      3. Ventricular wall weakness following myocardial infarction
      4. Ventricular dilatation in chronic congestive failure
      5. Inflammatory: bacterial, spirochaetal, viral, sarcoidosis, giant-cell myocarditis, Chagas' disease
      6. Metabolic: Marfan's syndrome, Hurler's syndrome, amyloidosis (Case 5, Fig. 5)
      7. Abnormally functioning papillary muscle (e.g., in hypertrophic subaortic stenosis)
      8. Trauma to papillary muscle
      9. Fibroelastosis and endomyocardial fibrosis
   b. *Chordal dysfunction*
      1. Congenital abnormalities (e.g., hypoplastic left heart syndrome, isolated congenital mitral insufficiency due to anomalous chordae tendinae)
      2. Spontaneous rupture (Case 3, Fig. 3)
      3. Damage after subacute bacterial endocarditis (SBE) or RHD (Case 2, Fig. 2)
      4. Connective tissue disease, Marfan's syndrome, Ehlers-Danlos syndrome, Hurler's syndrome, carcinoid syndrome.

forceful left ventricular impulse in the VIth intercostal space just outside the midclavicular line. On auscultation, S_1 was normal, S_2 was normally split (0.04 sec.) and an S_3 was present at the apex. A grade 3/6 pansystolic murmur was heard maximally at the apex. This had a crescendo in late systole and was transmitted along the entire precordium to the left axilla and to the back.

Phonocardiography confirmed these findings and in addition revealed a faint S_4 which coincided with a small "a" wave in the apex cardiogram (Fig. 1). A chest x-ray showed moderate left atrial and left ventricular enlargment and pulmonary vascular congestion. The electrocardiogram was compatible with left ventricular enlargement and displayed ST and T wave changes of the primary type. There was no evidence of myocardial scarring.

Even in the absence of definite evidence of past myocardial infarction by history or by electrocardiographic findings, on the basis of the data available a diagnosis of papillary muscle dysfunction of the mitral valve with regurgitation can be entertained. In this case a left ventricular angiocardiogram did indeed demonstrate moderately severe mitral insufficiency. The patient improved with digitalis and diuretics, but the apical systolic murmur with late accentuation continued to persist. Cases of this type are fairly frequent, often but not necessarily with evidence for past myocardial infarction.

CASE 2. Rupture of Chordae Tendineae of Mitral Valve Following Subacute Bacterial Endocarditis (Table 1, IIIb3).

S.K., a 45-year-old woman with known rheumatic heart disease involving the mitral valve. Several years ago the patient had two well documented episodes of subacute bacterial endocarditis which responded to treatment with cephalothin. She was admitted in marked congestive cardiac failure of recent origin. Examination revealed a markedly elevated jugular venous pressure with prominent "v" waves. There was hepatomegaly and pedal edema. The radial pulse was rapid and irregular. Blood pressure was 130 mm. Hg systolic, 70 mm. diastolic. There was a prominent and diffuse left ventricular impulse in the VIth left intercostal space in the anterior axillary line. On auscultation S_1 was loud, S_2 was narrowly split, and a loud S_3 gallop was heard at the apex. A grade 3/6 systolic murmur with late accentuation was present and it was best heard at the apex. An opening snap was heard along the left sternal border and the base of the heart, with an esti-

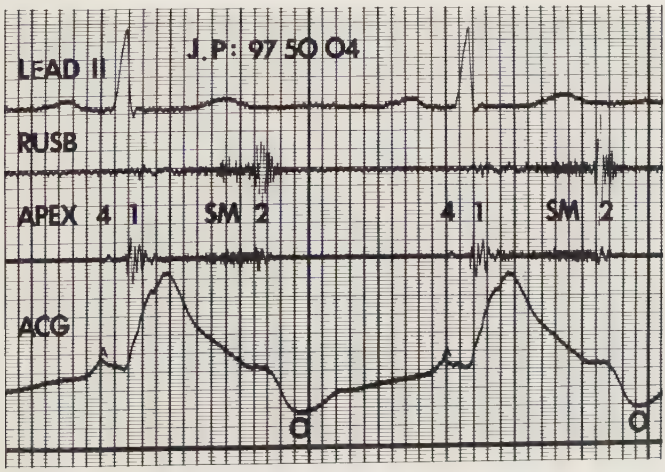


Figure 1. *Case 1: Papillary muscle dysfunction.*
Reading from above downward: lead II, phonocardiogram, and apex impulse cardiogram. RUSB, Right upper sternal border; SM, systolic murmur; ACG, apex cardiogram; O, point in the apex cardiogram which coincides with the opening of the mitral valve. Note the pansystolic murmur with late systolic accentuation.

mated A_2-OS duration of 0.06 sec. (Fig. 2). A chest radiogram confirmed the presence of cardiac enlargement with pulmonary edema. The electrocardiogram revealed evidence for left ventricular and left atrial hypertrophy. Atrial fibrillation was present.

Rapid deterioration of a patient during or after an attack of subacute bacterial endocarditis and presenting with the above clinical findings is suspicious of a rupture of the chordae tendineae of the mitral valve. Indeed, in this case at cardiac surgery, an isolated rupture of the chordae tendineae of the posterior mitral valve leaflet was discovered. The margins of the valves were intact. In spite of the opening snap recorded, no evidence for mitral stenosis was found. The mitral valve was excised and a Starr-Edwards prosthetic valve was sewn in. The patient made an excellent and uneventful recovery.

CASE 3. Spontaneous Rupture of Chordae Tendineae of Mitral Valve (Table 1, IIIb2).

I.C., a 61-year-old woman, was admitted in a state of severe congestive cardiac failure and with evidence of hepatic dysfunction. She deteriorated rapidly and soon went into hepatic coma. There was no previous history of rheumatic fever or subacute bacterial endocarditis. On examination the patient was clearly in gross congestive cardiac failure and appeared very drowsy and somewhat jaundiced. The radial pulse was rapid but regular. The blood pressure was 110 mm. Hg systolic, 75 mm. diastolic. There was a hyperactive left ventricular impulse in the VIth left intercostal space in the anterior axillary line. An apical systolic thrill was present. On auscultation the first sound was of normal intensity. It was preceded by a fourth sound of moderate intensity. The second sound was narrowly split and followed by a loud third heart sound. A very loud (grade 4/6) pansystolic murmur was present, rising to a crescendo in midsystole and best heard at the mitral area. It radiated along the left sternal border and aortic area, but was maximal at the apex (Fig. 3). In view of her rapid deterioration, cardiac surgery was thought to be indicated in spite of the obviously poor risk involved. A rupture of one of the chordae of the posterior cusp was discovered. The valve was replaced with a Starr-Edwards prosthesis. However, the postoperative course was stormy

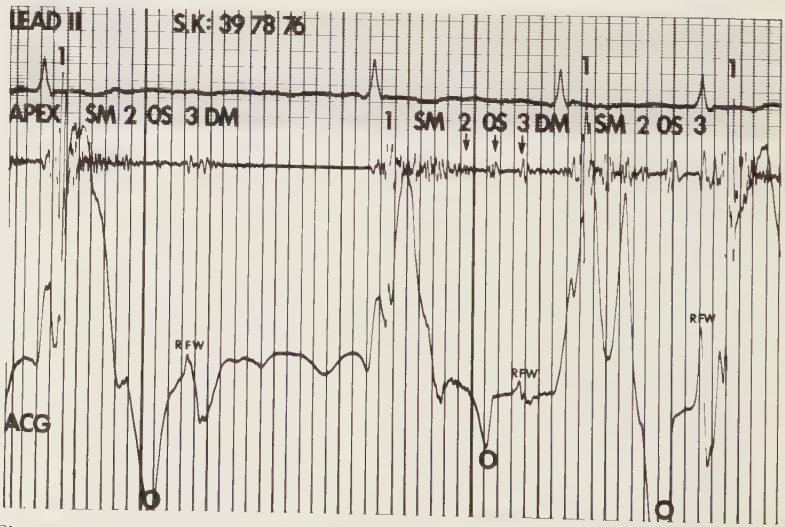


Figure 2. Mitral insufficiency due to *rupture of chordae tendineae* of the mitral valve following subacute bacterial endocarditis.

Note the opening snap (OS) of the anterior mitral cusp in pure mitral incompetence. A loud S_3 gallop is also present (case 2). Abbreviations are as in Figure 1.

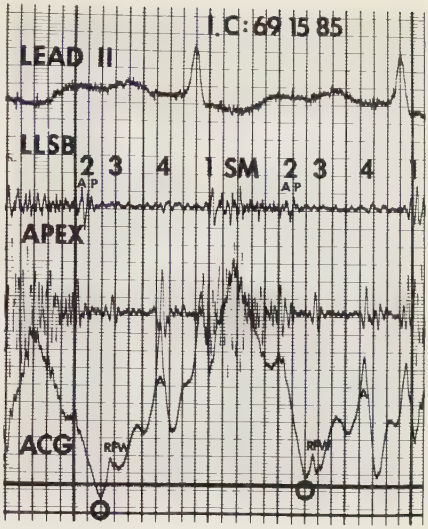


Figure 3. *Spontaneous rupture of a chorda of the mitral valve* with marked mitral insufficiency (case 3). The murmur is pansystolic in duration and rises to a crescendo in mid systole.

Abbreviations are as in Figure 1. A, Aortic closure; P, pulmonic closure; LLSB, left lower sternal border.

and the patient died on the following day due to uncontrollable myocardial failure. At autopsy there was no evidence whatever for rheumatic heart disease and the presumptive diagnosis of spontaneous rupture of a chorda tendinea was made. The liver showed marked passive venous congestion with areas of fibrosis.

CASE 4. Mitral Incompetence Due to a Fibrotic Mural Cusp (rheumatic heart disease?) (Table 1, Ia).

C.W., a 15-year-old boy, was found to have an apical systolic murmur in August 1965, during a routine physical examination. He was asymptomatic. During the following 2 years he gradually developed dyspnea on exertion and decreased exercise tolerance. There was no antecedent history of rheumatic fever, bacterial endocarditis, or trauma. While no abnormal findings other than the murmur were elicited before, physical examination now revealed a prominent jugular venous pulse and a radial pulse which was regular and rapid. The blood pressure was 110 mm. Hg systolic, 70 mm. diastolic. There was a hyperdynamic left ventricular impulse in the VIth intercostal space in the anterior axillary line. In addition, an apical systolic thrill was easily palpable. On auscultation a loud S_1 was elicited, which was followed by a high-pitched, somewhat squeaky pansystolic murmur ending in a normal S_2. S_2 was normally split. The loud first sound and the peculiar high-frequency apical murmur with an early peak suggested the possibility of a highly mobile anterior cusp of the mitral valve with mitral regurgitation (Fig. 4). In this patient, however, an opening snap, which is frequently seen in this condition, was not present. Cardiac catheterization and angiocardiography confirmed the presence of severe mitral regurgitation. At the time of cardiac surgery the aortic leaflet of the mitral valve was pliant and mobile, while the mural cusp was small and fibrotic. On pathological examination a diagnosis of rheumatic heart disease was made. The mitral valve was replaced with a Starr-Edwards prosthesis. The patient made an uneventful recovery.

This represents an example of an unusual deformation of a mitral valve with a mobile, well functioning second leaf—hence a loud first sound in addition to the murmur. The stiff, fibrotic valve is likely to be responsible for vibrations giving rise to a "musical" or high-pitched sound. Again, even an opening snap—considered the hallmark of mitral stenosis—can at times be heard under these conditions.

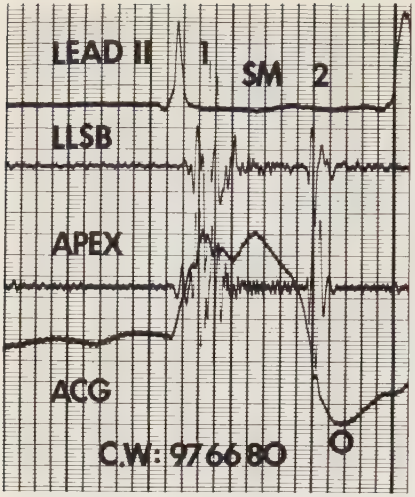


Figure 4. *"Floppy" valve syndrome* (fibrotic neural cusp).

Phonocardiogram and apexcardiogram from a patient with rheumatic mitral regurgitation (case 4). The anterior cusp was "floppy." Note the very loud S_1 which is followed by a pansystolic murmur.

CASE 5. Calcification of the Mitral Valve Annulus (Table 1, IIb and IIIa6).

S.D., a 19-year-old woman, presented with a history of a heart murmur since birth. Her major complaint was that of recurrent episodes of rapid heart action. Her tall stature, arachnodactyly, and arched palate gave the appearance of a patient with Marfan's syndrome. The pulse was regular. The blood pressure was 110 mm. Hg systolic, 60 mm. diastolic. There was a hyperdynamic left ventricular impulse in the Vth left intercostal space in the midaxillary line. There was also a moderately severe degree of pectus excavatum. The first sound was of normal intensity. It was followed by a holosystolic murmur with late systolic accentuation. The second sound was normally split, pulmonic closure being louder than aortic closure. A chest radiogram showed extensive calcification of the mitral annulus with left ventricular enlargement (Fig. 5). The electrocardiogram showed right axis deviation (+100°) and deep S waves in leads V_1 and V_2. A severe degree of mitral regurgitation was demonstrated by cineangiography.

In this example the cause for mitral regurgitation based on a calcified immobile valve ring remains obscure. One is tempted to relate the disorder to the genetic mesenchymal defect present in this patient. The possibility that the valve leakage was in part due to the redundant length of a chorda tendinea must be considered.

CASE 6. Mitral Incompetence Due to a Billowing Posterior Valve Cusp (Table 1, Id).

C.D. is a 16-year-old girl who visited the clinics "because of a heart murmur." She had no symptoms of heart disease. There was no history of rheumatic fever. The examination revealed few abnormalities; the jugular venous pulse was normal. The ventricular impulse was strong, but the apex beat was in the Vth left intercostal space in the midclavicular line. On auscultation S_1 was easily heard and was succeeded by a prominent midsystolic click. S_2 was normally split (Fig. 6). A loud systolic murmur (grade 4/6) introduced by the click was present in the supine position and best heard at the apex. The ECG revealed abnormal T waves in leads II, III, aVF, V_5, and V_6, suggesting left ventricular enlargement ("diastolic overload"). Left ventricular cineangiocardiography demonstrated mild mitral regurgitation. An abnormal billowing of the posterior mitral leaflet in systole was easily demonstrated.

CASE 7. 'Mitral' Incompetence in a Case of Corrected Transposition of the Great Vessels (Table 1, Ic).

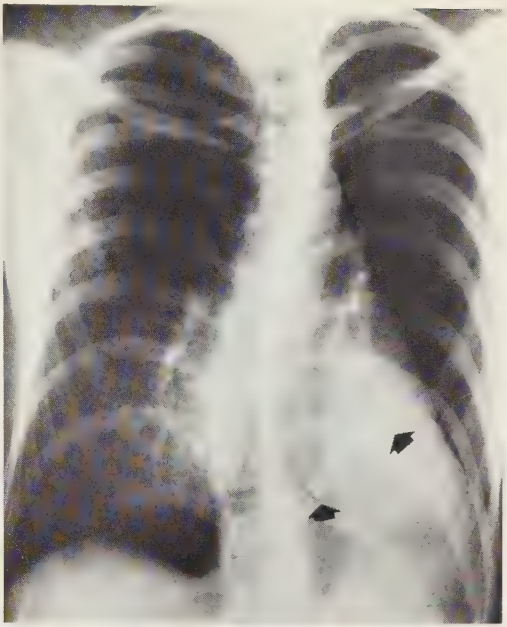


Figure 5. Calcification of the mitral annulus in a 19 year old patient with Marfan's syndrome (case 5). There is also a fair degree of pectus excavatum. Arrows point to calcified mitral annulus.

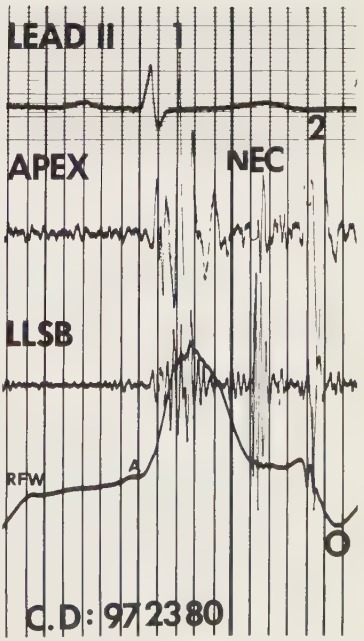


Figure 6. Lead II, phonocardiogram, and apex cardiogram from a patient who had a billowing posterior cusp of the mitral valve. The first sound is of unusual intensity and is succeeded by a prominent nonejection click (NEC). A low-frequency murmur follows the click (case 6).

A.S. was a 40-year-old woman who had been seen since 1959 with a presumptive diagnosis of rheumatic heart disease with mitral insufficiency. A systolic apical murmur was first detected at the age of 5. Two episodes of chest pain, fever, orthopnea and migratory polyarthritis at the ages of 12 and 28 were strongly suggestive of rheumatic fever. In 1959 she appeared chronically ill with a history of increasing fatigue, dyspnea on exertion, chest pain, and swelling of the ankle. At that time her jugular veins were distended, the radial pulse was 56 per minute and irregular, and the blood pressure was 135 mm. Hg systolic, 80 mm. diastolic. The point of maximal impulse was in the VIth left intercostal space, with a dynamic left ventricular beat extending to the anterior axillary line. A systolic apical thrill was present. On auscultation S_1 was muffled and followed by a grade 4 pansystolic murmur at the apex. The murmur radiated into the left axilla. S_2 was normally split and P_2 was louder than A_2 in the IInd left intercostal space. A grade 2 rumbling middiastolic murmur was also evident at the apex.

The ECG revealed atrial fibrillation with notched biphasic QRS complexes, in all limb leads, depressed ST segments and diphasic T waves in leads I, II, AVL, and V_4 through V_6. There was no transition of QRS across the precordium and in consequence no Q waves in V_5 or V_6 (Fig. 7). The appearance of the ECG was puzzling, since it was considered atypical for an example of rheumatic valvular disease. Chest radiograms confirmed the presence of cardiomegaly with left ventricular and bi-atrial enlargement (Fig. 8). Significantly, the right pulmonary artery was large, but the main left pulmonary artery was not well seen. The aortic arch appeared to be small.

Cardiac catheterization confirmed the diagnosis of mitral insufficiency with passive pulmonary hypertension. Mitral valvuloplasty was carried out by open heart surgery involving cardiopulmonary bypass. The patient had fairly good symptomatic relief following this procedure, but within 2 years her symptoms returned. She was managed medically from 1961 to 1966. During this period she had numerous episodes of atrial fibrillation which were successfully treated by medical means.

In March 1967 she was again operated upon and a Starr-Edwards prosthesis was inserted after removal of the mitral valve, which now was found to resemble an Ebstein anomaly. At the time of surgery the great vessels were found to be transposed and the diagnosis of corrected transposition was made. The patient made a good recovery initially. Unfortunately, she died a short while afterward from progressive myocardial failure. Autopsy confirmed the diagnosis of corrected transposition of the great vessels with malformation of the left A-V valve. In addition there was an anomalous left coronary artery arising from the pulmonary

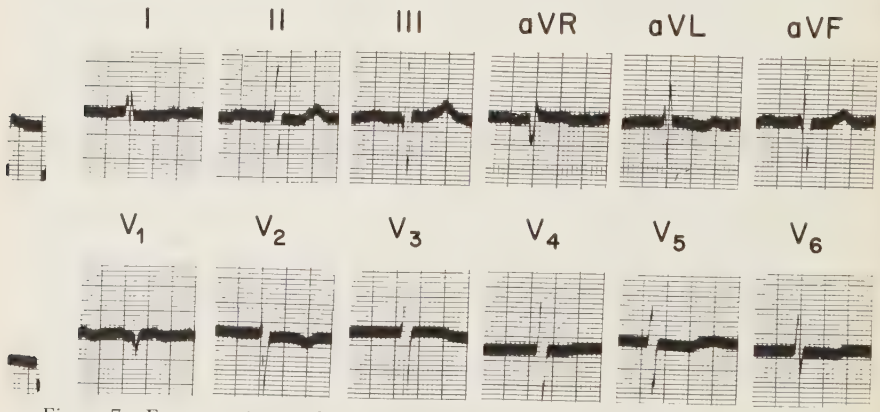


Figure 7. From a patient with corrected transposition of the great vessels (case 7). Note the absence of r in V_1, the leftward and superior orientation of the septal forces. Atrial fibrillation is present.

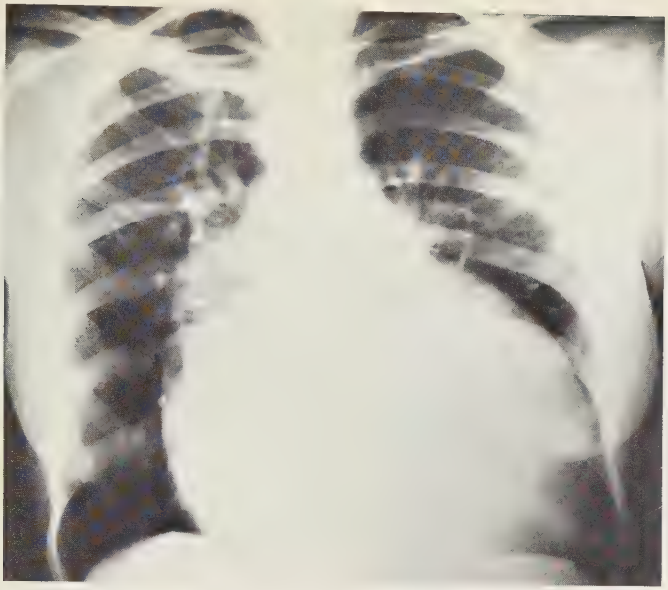


Figure 8. Chest radiogram, case 7. The left cardiac border is formed by the aorta. The pulmonary salient is absent because of its medial position in this anomaly, which was diagnosed as corrected transposition of the great vessels.

artery. The absence of septal forces in electrocardiographic leads V_5 and V_6 and the general direction of the septal forces (leftward and superiorly) should perhaps have suggested the diagnosis of corrected transposition of the great vessels. The enlarged heart, however, made such a prediction hazardous. Congenital heart disease with "mitral" (morphologically tricuspid!) valve anomaly is no contraindication to valve surgery!

Case 8. Anomalous Left Coronary Artery With Papillary Muscle Fibrosis (Table 1, IIIa2).

W.S. was seen at the age of 3 months at a local hospital where he was admitted in cardiac failure. At the age of 11 months he was evaluated by cardiac catheterization and angiocardiography. Left ventricular enlargement was demonstrated, but no intracardiac shunts suggestive of congenital cardiac defects were detected. The patient recovered. At the age of 3 years he began to complain of chest pain on exertion. He was not seen until the age of 9, when he was admitted with complaints of chest pain and episodes of unexplained fainting. On examination he again displayed the signs of congestive cardiac failure. Atrial fibrillation was present. On auscultation S_1 was soft and was immediately followed by a grade 2/6 pansystolic apical murmur. S_2 was normally split and pulmonary closure was louder than aortic closure. No gallop sounds were heard and no friction rubs were present. The ECG demonstrated left axis deviation on the frontal plane with inversion of T in I and V_2 but not in V_5 or V_6. There were no pathologic Q waves (Fig. 9). The chest x-ray was compatible with left ventricular and left atrial enlargement. Severe (rheumatic) mitral regurgitation was thought to be present and the mitral valve was replaced with a Starr-Edwards prosthesis, but the child died a few hours later. On autopsy the presence of an anomalous left coronary artery arising from the main pulmonary artery was discovered. The left posterior papillary muscle was atrophic.

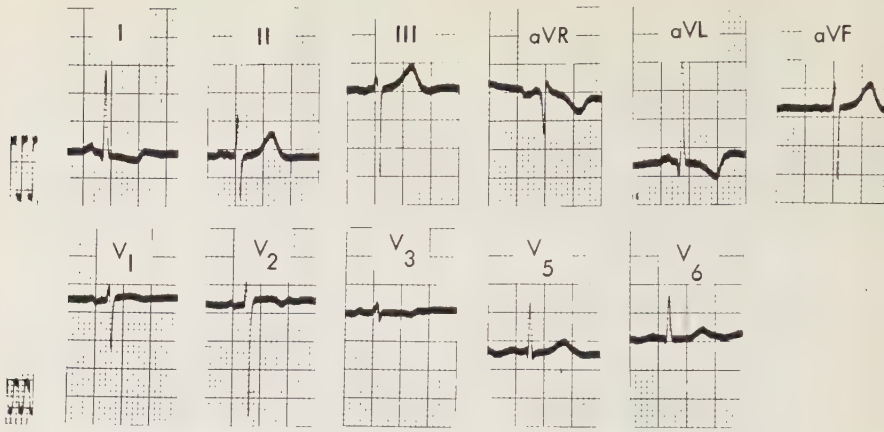


Figure 9. Electrocardiogram from a child with anomalous left coronary artery with *papillary muscle fibrosis*. Note left axis deviation in a young child with T wave inversion in lead I and aVL. However, no pathologic Q waves are present (case 8). There was no clinical or radiological evidence for ventricular enlargement.

DISCUSSION

The mitral valve apparatus is a complex structure. Normal physiological competence of the valve depends not only upon the valve leaflets but also on the delicately balanced action of the subvalvar structures (Fig. 10). The examples of mitral incompetence presented above represent different causes affecting one or more of these structures. Until recently several of these subvalvar forms of mitral incompetence (SMI) were considered rare, but lately the presence of SMI, particularly if due to papillary muscle dysfunction, has been more frequently recognized. While it occurs mostly in elderly subjects with ischemic heart disease, examples given above show that SMI may occur in almost any age group.

The anatomy of the mitral valve complex has been well reviewed recently by Silverman and Hurst.[31] It is perhaps worth recapitulating here that each papillary muscle supplies chordae tendineae to *both* mitral leaflets, so that rupture of a papillary muscle can affect the competence of both valves. Another point of note is that the anterolateral papillary muscle is well supplied with blood from the marginal branches of the circumflex and the diagonal branches of the anterior descending artery; the posteromedial papillary muscle, on the other hand, is less well supplied. It receives its nourishment from the right coronary artery and the posterior wall branches of the left circumflex artery. The posterior papillary muscle is therefore more vulnerable to ischemic damage.[7] Normal mitral valve function depends on the anatomic and mechanical integrity of the valvar structures, and the papillary muscles, as well as upon proper integration of the time relations between contraction of the papillary muscle and contraction of the free left ventricular wall. The function of papillary muscles to restrain the mitral valve is obvious.

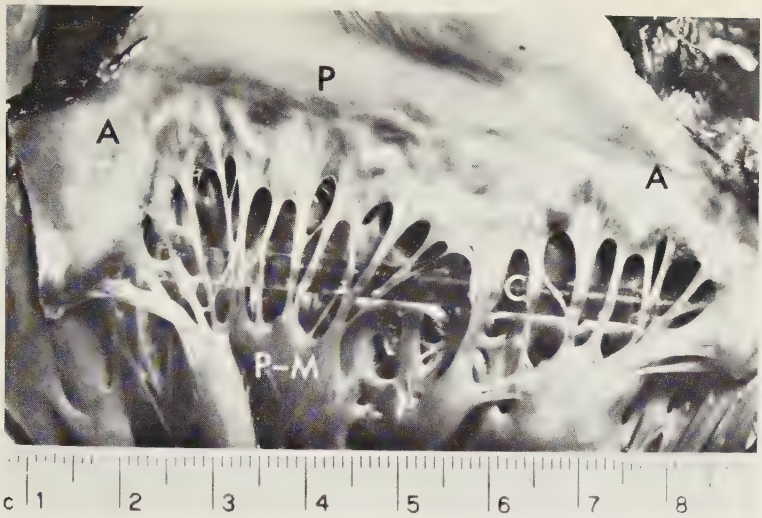


Figure 10. Normal mitral valve. The anterior cusp has been divided to show the subvalvar structures in detail. A, Anterior cusp; P, posterior cusp; C, chordae tendineae; A-L, anterolateral papillary muscle; P-M, posteromedial papillary muscle. Mitral insufficiency can result from the malfunction of any of these structures or from structural weakness of the myocardium supporting the base of the papillary muscles. (Preparation by Dr. M. Yacoub).

The syndrome of *papillary muscle dysfunction* as a clinical entity was described by Burch, DePasquale, and Phillips in 1963.[2] Failure of one or both papillary muscles to shorten during the ejection phase of ventricular systole, fibrosis and atrophy of a papillary muscle, or abnormal movements of the papillary muscles due to left ventricular dilatation, result in mitral incompetence. "Papillary muscle dysfunction," however, has a much wider significance for valvar regurgitation than its now somewhat overemphasized and overdiagnosed role in coronary artery disease would indicate. Depending upon the etiology of the papillary muscle dysfunction, apical systolic murmurs of varying characteristics may be heard. In general, a noncontracting papillary muscle in a heart of normal size is associated with a murmur which is late in onset and crescendo-decrescendo in quality, whereas in the dilated heart the murmur begins with the first sound, often engulfs this sound, and is usually pansystolic but may be decrescendo, plateau, or crescendo-decrescendo in quality.

Since myocardial infarction is the etiology of most cases of papillary muscle dysfunction, it can be seen how the additional burden of mitral incompetence can further compromise the function of the heart. It should be emphasized that in many of these cases the fundamental cause for mitral insufficiency is the weakness of the myocardium itself supporting the papillary muscles.

The clinical and auscultatory picture of this condition may be simulated by rupture of the interventricular septum, a complication of myocardial infarction. In some cases cardiac catheterization may be necessary to make the distinction between the two.

In acute subvalvar mitral incompetence due to *rupture of a papil-*

lary muscle or a chorda, a pansystolic murmur is the rule, as was mentioned by Raftery et al. in their review of the subject.[28]

At times the systolic regurgitation tide to the left atrium following acute papillary muscle rupture may be so pronounced that the "v" wave peak of the regurgitant stream from the left ventricle exceeds the pulmonary arterial pressure, with consequent reflux of oxygenated blood across the pulmonary artery.[33] If this occurs a left-to-right shunt may be diagnosed when blood samples are withdrawn for analysis. While rare, such events emphasize the extremely high pressures that can be produced within the left atrium when subvalvar mitral incompetence suddenly develops. The 'unprepared' and noncompliant left atrium is unable to absorb the brunt of the pressure head and consequently high pulmonary venous pressures leading to pulmonary edema can develop easily.

Rupture of the chordae tendineae usually produces a pansystolic murmur. With a flail anterior leaflet the murmur is best heard in the axilla and the paravertebral region.[10] In some cases a thrill may be felt along the spine. With chordal rupture of the posterior leaflet the murmur is heard toward the base of the heart and may radiate to the neck and head[27] (see Case 2).

Of considerable interest has been the description of the *'auscultatory-electrocardiographic' syndrome* of Humphries and McKusick[14] and Barlow and Bosman.[1] Case 6 in our series is illustrative of this condition. These patients demonstrate: (a) an aneurysmal protrusion of the posterior leaflet of the mitral valve; (b) a late systolic murmur of mitral incompetence; (c) a mid or late systolic click (nonejection click); and (d) diphasic or inverted T waves in leads II and III and peaked T waves in mid-precordial leads. Recently the familial nature of this syndrome has been pointed out.[30] It should be contrasted with the cases of mitral insufficiency recently described in *male* members of a family. Presumably these cases have an X-linked inheritance of mitral valvar disease.[23]

A mid or late systolic click at the apex in conjunction with an "innocent looking" late systolic murmur, especially in a young woman, should draw one's attention to this syndrome. Since the occurrence of subacute bacterial endocarditis has been reported in these subjects,[17] prophylaxis might have to be considered. Instances of arrhythmia and even sudden death have occurred.[12] Thus every late systolic murmur need not be innocent, as was once taught by Evans.[8] Particularly, the presence of the systolic click and electrocardiographic anomalies with or without characteristic murmurs should lead to careful evaluation.

In some cases it may be the anterior cusp that is "floppy" (see Case 4). A very loud S_1 preceding a pansystolic murmur should suggest the diagnosis. The etiology is obscure, and in our example rheumatic heart disease was diagnosed by the pathologist. Since the opening snap can occur, a clinical error is likely to be made as well. A loud opening snap as a sign of mitral incompetence with or without a ventricular filling sound has been described,[24, 25] particularly in mitral incompetence caused by disease of the posterior cusp.

Either mitral cusp may become floppy due to *redundant lengths of chordae tendineae*, as in Marfan's syndrome[29] or Ehlers-Danlos dis-

ease.[20] Sudden tensing of the slack chordae during systole may produce a snapping sound (chordal snap). In addition the valve cusps themselves may undergo a curious myxomatous transformation with eventual calcification. Case 5 in our series probably belongs to this group.

Chordal rupture may be either spontaneous (Case 3) or due to bacterial endocarditis (Case 2). In both instances the diagnosis may at times be difficult. Sudden severe deterioration in ventricular function accompanied by a loud left atrial gallop with a left atrium of normal size and a pansystolic crescendo-decrescendo murmur at the apex should suggest the possibility of a ruptured chorda. Because of the poor prognosis, prompt surgical measures are mandatory in these cases.[15]

Bad alignment of the papillary muscles may also occur in idiopathic hypertrophic subaortic stenosis. The asymmetrically hypertrophied musculature of the left ventricle may distort the papillary muscles, particularly those guiding the anterior cusp. Furthermore, abnormal contraction of the papillary muscle may cause systolic reopening of the mitral valve with consequent insufficiency.[4] Leaflet mobility may be restricted in anomalous mitral arcade.[16] The papillary muscles may also be abnormally inserted in an unusual location in the ventricle.[6] Finally the whole valve apparature may be poorly developed, or an Ebstein-like malformation in a functionally corrected transposition may be present[21] (Case 7).

The presence of mitral regurgitation in an infant or child who presents with episodes of pallor, cyanosis, colicky pain, and tachypnea should suggest the presence of an anomalous left coronary artery with ischemic fibrosis of a papillary muscle.[26] Case 8 illustrates this uncommon but interesting condition in a patient who survived until 9 years of age.

Rarely a space-occupying lesion in the left atrium, such as myxoma, may interfere with mitral valve function. While the clinical features usually resemble those of mitral stenosis, occasionally mitral regurgitation is simulated.[22]

The purpose of this review was not to describe each form of mitral incompetence in detail but to indicate certain features of nonrheumatic disease of the mitral valve system and to point to the deceptive nature of the so-called innocent late systolic murmur. It is of particular interest in nearly all of the cases of subvalvar mitral incompetence that the first sound is either normal or increased in intensity. This contrasts with the situation in rheumatic valvar incompetence, where S_1 is often diminished in intensity or engulfed in the pansystolic murmur. Another diagnostic clue is the presence of a mid to late ejection click, which heralds a late systolic murmur. The presence of an opening snap is a sign of peculiar valve movement and is not strictly a sign of mitral stenosis. A proper appreciation of mitral incompetence demands an interpretation of the various physical findings in terms of disorders of the various components of the mitral valve complex.

REFERENCES

1. Barlow, J. B., and Bosman, C. K.: Aneurysmal protrusion of the posterior leaflet of the mitral valve: An auscultatory-electrocardiographic syndrome. Amer. Heart J., *71*:166, 1966.

2. Burch, G. E., DePasquale, N. P., and Phillips, J. H.: Clinical manifestations of papillary muscle dysfunction. Arch. Int. Med., *112*:112, 1963.
3. Cabot, R. C.: Facts on the Heart. Philadelphia, W. B. Saunders, Co., 1926.
4. Cohen, J., Effat, H., Goodwin, J. R., Oakley, C. M., and Steiner, R. E.: Hypertrophic obstructive cardiomyopathy. Brit. Heart J., *26*:16, 1964.
5. Corvisart, J. N.: Essai sur les maladies et les lesions organiques du Coeur et des gros Vaisseaux. 1806. (English edition: An Essay on the Organic Diseases and Lesions of the Heart and Great Vessels. Translated from the French with notes by Jacob Gates, Boston, 1812. Reprinted in paper covers by Hafner, New York, in 1962.)
6. Edwards, J. E.: The problem of mitral insufficiency caused by accesory chordae tendinae in persistent common atrioventricular canal. Proc. Mayo Clin., *35*:299, 1960.
7. Estes, E. H., Jr., Dalton, F. M., Entman, M. D., Dixon, H. B. II, and Hackel, D. B.: The anatomy and blood supply of the papillary muscles of the left ventricle. Amer. Heart J., *71*:356, 1966.
8. Evans, W.: Cardiography. London, Butterworth, 1948.
9. Freeman, A. R., and Levine, S. A.: The clinical significance of the systolic murmur. Ann. Int. Med., *6*:1371, 1933.
10. Giuliani, E. R.: Mitral valve incompetence due to flail anterior leaflet. Amer. J. Cardiol., *20*:784, 1967.
11. Griffith, J. P. C.: Mid-systolic and late systolic mitral murmur. Amer. J. Med. Sci., *104*: 285, 1892.
12. Hancock, E. W., and Cohn, K.: The syndrome associated with mid-systolic click and late systolic murmur. Amer. J. Med., *41*:183, 1966.
13. Hope, J.: A treatise on diseases of the heart and great vessels. London, William Kidd, 1832.
14. Humphries, J. O., and McKusick, V. A.: The differentiation of organic and "innocent" systolic murmurs. Progr. Cardiovasc. Dis., *5*:152, 1962.
15. Klughaupt, M., Flamm, M., Hancock, W. E., and Harrison, D.: Nonrheumatic mitral insufficiency. Circulation, *39*:307, 1969.
16. Layman, T. E., and Edwards, J. E.: Anomalous mitral arcade. Circulation, *35*:389, 1967.
17. LeBauer, E. J., Perloff, J. K., and Kelliher, T. F.: Isolated systolic click with bacterial endocarditis. Amer. Heart J., *73*:534, 1967.
18. Lewis, T.: Diseases of the Heart. London, MacMillan, 1933.
19. MacKenzie, J.: Principles of Diagnosis and Treatment of Heart Affections. 2nd ed. London, Oxford Medical Publications, 1916.
20. Madison, W. M., Bradley, E. J., and Castillo, A. J.: Ehlers-Danlos syndrome with cardiac involvement. Amer. J. Cardiol., *11*:689, 1963.
21. Malers, F., Björk, V. O., Cullhed, I., and Lodin, H.: Transposition functionally totally corrected, associated with mitral insufficiency. Amer. Heart J., *59*:816, 1960.
22. Marpole, D. G. F., Kloster, F. E., Bristow, D. J., and Griswold, H. E.: Atrial myxoma, a continuing diagnostic challenge. Amer. J. Cardiol., *23*:597, 1969.
23. Monteleone, P. M., and Fajan, L. F.: Possible X-linked congenital heart disease. Circulation, *39*:611, 1969.
24. Nixon, P. G. F.: The third heart sound in mitral regurgitation. Brit. Heart. J., *23*:677, 1961.
25. Nixon, P. G. F., Wooler, G. H., and Radigan, L. R.: The opening snap in mitral incompetence. Brit. Heart J., *22*:395, 1960.
26. Noren, G. R., Raghib, G., Moller, J. H., Amplatz, K., Adams, P., and Edwards, J. E.: Anomalous origin of the left coronary artery from the pulmonary trunk with special reference to the occurrence of mitral insufficiency. Circulation, *30*:171, 1964.
27. Osmundson, P. J., Callahan, J. A., and Edwards, J. E.: Mitral insufficiency from ruptured chordae tendineae simulating aortic stenosis. Proc. Staff Meet. Mayo Clin., *33*:235, 1958.
28. Raftery, E. B., Oakley, C. M., and Goodwin, J. F.: Acute subvalvar mitral incompetence. Lancet, *2*:360, 1966.
29. Segal, B., Kasparian, H., and Likoff, W.: Mitral regurgitation in a patient with the Marfan syndrome. Dis. Chest., *41*:457, 1962.
30. Shell, W. E., Walton, J. A., Clifford, M. E., and Willis, P. W.: The familial occurrence of the syndrome of mid-late systolic click and late systolic murmur. Circulation, *39*:327, 1969.
31. Silverman, M. E., and Hurst, J. W.: The mitral complex. Amer. Heart J., *76*:399, 1968.
32. Sprague, H. D., and White, P. D.: A comparative study of rheumatic mitral regurgitation and stenosis. Amer. Heart J., *1*:629, 1926.
33. Tatooles, G. J., Gault, J. G., Mason, D. T., and Ross, J., Jr.: Reflux of oxygenated blood into the pulmonary artery in severe mitral regurgitation. Amer. Heart J., *75*:102, 1968.
34. Vieussens, R.: Traité nouveau de la structure et des causes du mouvement naturel du Coeur. Toulouse, Jean Guillemelte, 1715.
35. Wood, P.: An appreciation of mitral stenosis. Brit. Med. J., *1*:1051, 1113, 1954.

950 East 59th Street
Chicago, Illinois 60637 (Dr. Hecht)

Index

NOTE: Page numbers of symposium and article titles are in **boldface** type.

ACID-PEPTIC disease, hyperparathyroidism and, 132
Adrenocorticosteroids. See *Corticosteroids.*
ALG. See *Antilymphocyte globulin.*
Allergy, to antibiotics, **59-73**
  alternate drugs and, 69
  corticosteroids and antihistamines for, 71
  desensitization and, 70
  incidence of, 62-64
  mechanism of, 64-66
  prevention of reactions, 66-68
  treatment of reactions, 68-69
ALS. See *Antilymphocyte serum.*
Amenorrhea, oral contraception and, 167
Aminoglycosides, allergic reactions to, incidence of, 63
Anaphylactoid reactions to antibiotics, 61, 62
Anemia, erythropoietin and, 175-177
Antibiotics, allergy to, **59-73**
  alternate drugs and, 69
  corticosteroids and antihistamines for, 71
  desensitization and, 70
  incidence of, 62-64
  mechanism of, 64-66
  prevention of reactions, 66-68
  treatment of reactions, 68-69
  alternate, history of allergic reaction and, 69
  contaminants of, reactions and, 64
Antigens, transplantation, 45-53
  detection of, 46-52
  immunologic tolerance and, 54-55
  inheritance of, 52-53
Antihistamines, for allergy to antibiotics, 71
Antilymphocyte globulin, after kidney transplants, 87
Antilymphocyte serum, after kidney transplants, 87
Ascorbic acid, deficiency, osteoporosis and, 145
Autoimmunity, exophthalmic goiter and, 120-123
  skin disease and, **95-105**
  thyroid disease and, 120
Azathioprine, after kidney transplants, 84

BLOOD, cells, differentiation of, erythropoietin and, 174
  coagulation. See *Coagulation.*
  pressure, central venous, in myocardial infarction, 238-239
Bone, changes, in hyperparathyroidism, 134-135
  metabolism of, osteoporosis and, **141-152**
Bullous pemphigoid, autoimmunity and, 100-104

CALCIUM, metabolism, osteoporosis and, 145-146
  serum, hyperparathyroidism and, 132
  urinary, hyperparathyroidism and, 135
Carbohydrates, metabolism of, oral contraception and, 166
Cephalosporins, allergic reactions to, incidence of, 62
Chloramphenicol, allergic reactions to, incidence of, 63
Chromaffin tissue, glucagon and, 186
Coagulation, oral contraception and, 167
Computers in medicine, Regional Medical Programs and, 25
Contraception, oral, current status of, **163-171**
Coronary care units, nurses and physicians for, Regional Medical Programs and, 23-24
Corticosteroids, after kidney transplants, 87
  for allergy to antibiotics, 71
  for ophthalmologic complications of exophthalmic goiter, 126

DESENSITIZATION, for allergy to antibiotics, 70
Digitalis, cellular action of, **221-234**
  for shock in myocardial infarction, 241
Donor selection. See under *Transplantation.*
Dopamine, for shock in myocardial infarction, 241
Drugs, immunosuppressive, after kidney transplants, 84
Dyslipoproteinemias, diagnosis of, **153-159**

EDUCATION, continuing, current concepts and future needs, **5-18**
Electrophoresis, in study of plasma lipoproteins, 153
Endocrine glands, secretions of, osteoporosis and, 147-149
Erythromycin, allergic reactions to, incidence of, 63
Erythropoietin, current status of, **173-181**

FEVER, as reaction to antibiotics, 61
Fluoride, metabolism, osteoporosis and, 146-147

GENTAMICIN, allergic reactions to, incidence of, 63
Globulin, antilymphocyte (ALG). See *Antilymphocyte globulin.*
Glucagon, physiologic role and therapeutic use, **183-190**
Glucocorticoids, osteoporosis and, 147-148
Glycosides, cellular action of, **221-234**
Goiter, exophthalmic, autoimmunity and, 120-123
  choice of treatment for, 123-126
  ophthalmologic complications, 126-127
Goodpasture's syndrome, autoimmunity and, 100
Graft vs. host assay, in tissue typing, 46
Grafts, rejection, management of, 88
  mechanism of, 53-54
Graves' disease. See *Goiter, exophthalmic.*

HASHIMOTO'S thyroiditis, autoimmunity and, 121
Heart, denervated, function of, 109-110
  electrical slowing of, **247-259**
  failure, congestive, hypertrophy and, 218
  glucagon and, 186
  hypertrophy, cellular and subcellular basis of, **211-220**
  pacemakers, paired ventricular stimulation by, **247-259**
  transplantation of, hemodynamic observations after, **107-113**
  manifestations of rejection after, 110-111
  valves, incompetence, **261-274**
Hematopoiesis, disturbed, as reaction to antibiotics, 61
Histocompatibility antigens. See *Antigens, transplantation.*
HL-A locus, 52
Hormones. See under name of hormone.
Hypercalcemia, hyperparathyroidism and, 132
Hyperlipoproteinemias, diagnosis of, 154-156
Hyperparathyroidism, primary, diagnosis of, simplified approach to, **131-139**
  treatment of, 137-138
Hypolipoproteinemias, diagnosis of, 156-157
Hypophosphatemia, hyperparathyroidism and, 134

IMMUNITY. See also *Autoimmunity.*
  tissue, biological significance of, 55
  tissue typing and, **43-58**
Immunologic tolerance, 54-55
Immunosuppression. See also *Drugs, immunosuppressive.*
  after kidney transplantation, 85-86
Infection, complicating kidney transplantation, 89
Insulin, biosynthesis of, 191-192
  osteoporosis and, 148
  protein synthesis and, **201-207**
Iodine, protein-bound, thyroid function and, factors in interpretation, 117
  radioactive, for exophthalmic goiter, 125
  uptake test, factors in interpretation, 118
Isoproterenol, for shock in myocardial infarction, 240

KANAMYCIN, allergic reactions to, incidence of, 63
Kidneys, systemic lupus erythematosus and, 96
  transplantation of, complications of, 88
  dialysis and, 75-77
  donor selection in, 79-83
  medical criteria for, 78
  patient selection and management in, **75-94**
  preparation for, 83
  postoperative care, 84
  tubular necrosis of, acute, after transplantation, 89

LIPOPROTEINS, plasma, abnormal, diagnosis of, **153-159**
Liver, function of, oral contraception and, 166
Lupus erythematosus, systemic, autoimmunity and, 95-100
  reactions to antibiotics resembling, 61
Lymphocyte transfer test, in tissue typing, 46
Lymphocytotoxicity test, in tissue typing, 47-50

MACROLIDE antibiotics, allergic reactions to, incidence of, 63
Medicine, modern challenges in, symposium on, **1-274**
Menstruation, disorders of, oral contraception and, 167
Metabolism, carbohydrate, oral contraception and, 166
  of bone, osteoporosis and, **141-152**
Methimazole, for exophthalmic goiter, 124
Mitral valve incompetence, **261-274**
Mixed lymphocyte culture, in tissue typing, 50-52
Myocardium, contraction, cellular events of, 224-225
  excitation, cellular events of, 222-224
  infarction, hemodynamic effects and results of therapy, **235-246**

NEOMYCIN, allergic reactions to, incidence of, 63
Neutral third party test, in tissue typing, 46
Norepinephrine, for shock in myocardial infarction, 240
Nutrition, osteoporosis and, 144-147

OSTEOPOROSIS, bone metabolism and, **141-152**

PACEMAKERS, cardiac, paired ventricular stimulation by, **247-259**
Parathormone. See *Parathyroid hormone.*
Parathyroid hormone, osteoporosis and, 149
radioimmunoassay, hyperparathyroidism and, 136-137
PBI. See *Iodine, protein-bound.*
Pemphigus vulgaris, autoimmunity and, 100-104
Penicillin, allergic reactions to, contaminants and, 64
incidence of, 62
prevention of, 66-68
Phosphate, tubular reabsorption of, hyperparathyroidism and, 135
Photosensitivity reaction to antibiotics, 60
Polycythemia, erythropoietin and, 177-178
Potassium iodide, for thyrotoxicosis, 127
Proinsulin, **191-199**
Propranolol, for thyrotoxicosis, 128
Propylthiouracil, for exophthalmic goiter, 124
Protein, synthesis of, cardiac hypertrophy and, 214-216
insulin and, **201-207**

REGIONAL Medical Programs, rural, **19-28**
urban, **29-33**
Ristocetin, allergic reactions to, incidence of, 64
RMP. See *Regional Medical Programs.*

SERUM, antilymphocyte (ALS). See *Antilymphocyte serum.*
Shock, myocardial infarction and, 237-238
Skin, disease, autoimmune factors in, **95-105**
osteoporosis and, 143
reactions to antibiotics, 60
SLE. See *Lupus erythematosus, systemic.*
Steroids, therapy. See *Corticosteroids.*
Streptomycin, allergic reactions to, incidence of, 63
Susceptibility factors. See *Antigens, transplantation.*

T_3 RESIN uptake test, factors in interpretation, 118
Tanned red cell agglutination test, for thyroid autoantibodies, 120
Test(s), for transplantation antigens, 46-52
tanned red cell agglutination, for thyroid autoantibodies, 120
Tetracyclines, allergic reactions to, incidence of, 63
Thromboembolism, oral contraception and, 168
Thyroid gland, autoantibodies, tanned red cell agglutination test for, 120
disorders of, autoimmunity and, 120
current concepts in management of, **117-130**
function of, oral contraception and, 166
tests of, factors in interpretation, 117-118
subtotal resection, for exophthalmic goiter, 126
Thyroid hormone, osteoporosis and, 148-149
Thyroxine, serum, measurement of, 118-119
Tissue typing, 46-52
and tissue immunity, **43-58**
"Transplant lung," 89
Transplantation, donor selection in, 79-81
early studies of, 44
genetic locus in man, 52
immunity and, 44-45
of heart, hemodynamic observations after, **107-113**
of kidneys, patient selection and management in, **75-94**
Transplantation antigens. See under *Antigens.*
Triacetyloleandomycin, allergic reactions to, incidence of, 63
Tumors, malignant, vs. hyperparathyroidism, 133

ULTRACENTRIFUGATION, in study of plasma lipoproteins, 153
University of Chicago medical school system, **35-40**
Uremia, chronic, dialysis vs. transplantation for, 77

VALVES, of heart. See under *Heart.*
Vancomycin, allergic reactions to, incidence of, 64
Vitamin C. See *Ascorbic acid.*

FIRST CLASS
PERMIT NO. 101
PHILADELPHIA, PA.

BUSINESS REPLY MAIL
NO POSTAGE STAMP NECESSARY IF MAILED IN UNITED STATES

POSTAGE WILL BE PAID BY

W. B. Saunders Company
West Washington Square
Philadelphia, Pa. 19105

On

DATE SLIP

Make sure
The Clinic
with you. I
attached c
to include
one month
to be proce

Gentleme

I am a subs

☐ THE M
☐ THE S
☐ THE F
☐ THE C
☐ THE I
☐ THE N

As of
new addres

New Addre

NAME

STREET

STATE

ZIP CODE

Old Address:

STREET CITY

STATE ZIP CODE

JAN 17 1970 JAN 17 1970
JAN 29 1970 JAN 29 1970
JUL 6 1970 JUL 11 1970
AUG 12 1970 AUG 12 1970
AUG 24 1970 AUG 21 1970
NOV 14 1970 NOV 13 1970
NOV 24 1970 NOV 25 1970
JAN 11 1971 JAN 13 1971
JAN 13 1971 JAN 16 1971
FEB 4 1971 FEB 3 1971
FEB 20 1971 FEB 21 1971
APR 9 1971 APR 6 1971
APR 9 1971 APR 10 1971
JUL 22 1971 JUL 22 1971
JAN 14 1972 JAN 12 1972
APR 2 1979 APR 2 1979

CORNELL UNIVERSITY
MEDICAL COLLEGE
1300 York Ave., N.Y.C

JAN 9 1970

MEDICAL LIBRARY

C.U.M.C.
C0114J